

STIC-Biotech/ChemLib

81976

From: Huff, Sheela
Sent: Tuesday, December 10, 2002 7:28 AM
To: STIC-Biotech/ChemLib
Subject: search request for 09/628126

RECEIVED

DEC 10 2002

Please search SEQ ID No. 22 and aa 19-390 of SEQ ID No. 2

FROM SHELL DIVISION
(STIC)

Thanks

Sheela Huff
Art Unit 1642
CMI-8807
mailbox 8E12
305-7866

Point of Contact
P. Sheppard
Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 12/17/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 16, 2002, 05:36:01 ; Search time 2219 Seconds
(without alignments)
4729.462 Million cell updates/sec

Title: US-09-628-126-22

Perfect score: 648

Sequence: 1 ATGCATGTGCGGCGGCTC.....TATACAGTAATTCACACTGA 648

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmo:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	646.8	99.8	910	9 AL583443	AL583443
2	381.4	58.9	487	12 BF890112	BF890112 289772 MA
3	290.8	44.9	491	13 BI774014	BI774014 465791 MA
4	283.6	43.8	318	12 BF082784	BF082784 PMI-BT075
5	275	42.4	670	17 AZ786450	AZ786450 2M0032A02
6	267.4	41.3	276	10 AW602310	AW602310 RC5-BT055

7	254.8	39.3	277	10	BE073509	BE073509 RC5-BT055
c	248.8	38.4	634	12	BG092633	BG092633 mac14f02.
9	240.6	37.1	462	13	BI774765	BI774765 466980 MA
10	188.4	29.1	550	10	AW654920	AW654920 105425 MA
11	151.6	23.4	602	10	BB619915	BB619915 BB619915
12	86.8	13.4	219	14	W21054	W21054 zb53e12.r1
13	86.8	13.4	524	9	AT734260	AT734260 zb53e12.y
14	86.8	13.4	552	17	AZ711040	AZ711040 RPCI-24-1
15	70	10.8	463	12	BG091233	BG091233 mac14f02.
16	56.8	8.8	621	9	AI982035	AI982035 pat.pk007
c	39.4	6.1	570	12	BE920566	BE920566 EST424335
18	39.4	6.1	624	10	AV858691	AV858691 AV858691
19	39.4	6.1	691	10	AV849264	AV849264 AV849264
20	38.4	5.9	567	13	BI394782	BI394782 pglp1n.pk0
21	38.2	5.9	655	17	BH095725	BH095725 RPCI-24-3
c	37	5.7	564	17	AZ426088	AZ426088 1M0206P22
23	36.8	5.7	262	9	AI633788	AI633788 tt28e08.x
c	36.8	5.7	262	12	BF924792	BF924792 IL5-NT022
c	36.8	5.7	290	10	AW593961	AW593961 hg33d03.x
c	36.8	5.7	336	14	BM695772	BM695772 UI-E-CL1-
c	36.8	5.7	350	9	AI041228	AI041228 ov56a11.x
c	36.8	5.7	368	9	AA662727	AA662727 nu91c07.s
c	36.8	5.7	383	13	BI496339	BI496339 df124607.
c	36.8	5.7	389	9	AI983451	AI983451 wt49c08.x
c	36.8	5.7	390	9	AI083658	AI083658 qf26h04.x
c	36.8	5.7	416	14	D56225	D56225 HUMA20D12B
c	36.8	5.7	436	9	AI815633	AI815633 au43d04.y
c	36.8	5.7	535	9	AI144728	AI144728 am61c12.x
c	36.8	5.7	601	13	BI669176	BI669176 603295413
c	36.8	5.7	603	10	BE254310	BE254310 601115042
c	36.8	5.7	607	13	BI491553	BI491553 df11c12.w
c	36.8	5.7	612	9	AI878898	AI878898 au51a10.y
c	36.8	5.7	625	12	BG703159	BG703159 602686051
c	36.8	5.7	635	12	BG708013	BG708013 602671338
c	36.8	5.7	677	14	BM676656	BM676656 UI-E-EJO-
c	36.8	5.7	703	12	BG699283	BG699283 602678955
c	36.8	5.7	710	12	BG699790	BG699790 602681496
c	36.8	5.7	710	13	BI550072	BI550072 603192540
c	36.8	5.7	711	13	BI545106	BI545106 603242708

ALIGNMENTS

RESULT 1
AL583443
LOCUS
DEFINITION AL583443 LTI_NFL010_BC2 Homo sapiens CDNA clone CS0DL012YA12 5
prime, mRNA sequence.
ACCESSION AL583443
VERSION AL583443.1 GI:12952410
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 910)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope, Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Location/Qualifiers
1..910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DL012YA12"
/clone_lib="LTI_NFL010_BC2"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com URL : http://fulllength.invitrogen.com"

BASE COUNT 244 a 232 c 204 g 225 t 5 others
ORIGIN

Query Match 99.8%; Score 646.8; DB 9; Length 910;
Best Local Similarity 99.5%; Pred. No. 4.9e-183;
Matches 645; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCATGCGCGGGGCTCGTGGCCAGCCACCTGGGACACACGCGCGAGCTATTTC 60
Db 245 ATGCATGCGCGGGGCTCGTGGCCAGCCACCTGGGACACACGCGCGAGCTATTTC 304
Qy 61 TATTGACACAGGCACCTCTGGCTGTGCGCTTGTCTTACCGTGGCCACTATTATGGTG 120
Db 305 TATTGACACAGGCACCTCTGGCTGTGCGCTTGTCTTACCGTGGCCACTATTATGGTG 364
Qy 121 TTGGTGTTCAGAGACGAGCTCCATCCCACTCACCTGACAGCTCCCTCCAAAGA 180
Db 365 TTGGTGTTCAGAGACGAGCTCCATCCCACTCACCTGACAGCTCCCTCCAAAGA 424
Qy 181 GGAATGCTCAGAGACCTCTTATGATCCCTGAAAGAGCTCCATCAAGAGTCATGG 240
Db 425 GGAATGCTCAGAGACCTCTTATGATCCCTGAAAGAGCTCCATCAAGAGTCATGG 484
Qy 241 GCCTACCTCCAAAGTGGCAAGCATCTAAACAAACCAAGTGTCTTGGAAACAAAGATGGC 300
Db 485 GCCTACCTCCAAAGTGGCAAGCATCTAAACAAACCAAGTGTCTTGGAAACAAAGATGGC 544
Qy 301 ATTCCTCCAGTGCAGATATCAGAGATGGGAATCTGGTGATCCAAATCCCTGGTTGTAC 360
Db 545 ATTCCTCCAGTGCAGATATCAGAGATGGGAATCTGGTGATCCAAATCCCTGGTTGTAC 604
Qy 361 TTCATCATTTGCCAAGTGCAGTTCTTGTACAAATGCCCAATATCTGTCTGATCTGAAG 420
Db 605 TTCATCATTTGCCAAGTGCAGTTCTTGTACAAATGCCCAATATCTGTCTGATCTGAAG 664
Qy 421 TTGAGCTTCTCATCAACAGCATATCAAAAACAGGCGCTGGTGACAGTGTGTGAGTCT 480
Db 665 TTGAGCTTCTCATCAACAGCATATCAAAAACAGGCGCTGGTGACAGTGTGTGAGTCT 724
Qy 481 GGAATGCAACGAAACAGTATACCAAGATCTCTCAATTTCTGCTGGATTAACCTGCAG 540
Db 725 GGAATGCAACGAAACAGTATACCAAGATCTCTCAATTTCTGCTGGATTAACCTGCAG 784
Qy 541 GTCACACCCACCATATCAGTCAATGTGGATACATCCAGTACATAGATACAGCACCTTT 600
Db 785 GTCACACCCACCATATCAGTCAATGTGGATACATCCAGTACATAGATACAGCACCTTT 844
Qy 601 CCTCTTGAGATGTGTGCTCCTCTTATACATTAATTCAGACTGA 648
Db 845 CCTCTTGAGATGTGTGCTCCTCTTATACAGTAATTCAGACTGA 892

RESULT 2
BF890112 LOCUS BF890112 487 bp mRNA linear EST 25-APR-2001
DEFINITION 289772 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BF890112
VERSION BF890112.1 GI:12281570
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 487)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett G.L., Heaton,M.P., Laegreid,W., Rohrer,G.A., Chitko-McKown,C.G., Perlea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGACG
Plate: 96 row: N column: 9
Seq primer: ATTAGTGACACTATAG.

FEATURES
source
1..487
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."
148 a 125 c 106 g 108 t

BASE COUNT 148 a 125 c 106 g 108 t
ORIGIN
Query Match 58.9%; Score 381.4; DB 12; Length 487;
Best Local Similarity 86.4%; Pred. No. 2.1e-103;
Matches 421; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
Qy 109 ACTATTATGGTGTGTGCTTTCAGAGGAGGACTTCCATTCCTCCCACTCCAGTCAACGTC 168
Db 1 ACCATTATGGTGTGTGCTTTCAGAGGAGGACTTCCATTCCTCCCACTCCAGTCAACGTC 60
Qy 169 CCCCTCAAGAGGAGAAATTCCTCAGAGACCTCTTATGTATCTGTAAGAGCTCCATTC 228
Db 61 CCCCTTAAAGGAGGAAATTCCTCAGAGGAGCATCTCATGTATCTGTAAGAGGCTCCATTC 120
Qy 229 AAGAAGTCATGGCCCTACCTCCCAAGTGGCAAGCATCTAAACAAACCAAGTCTCTGTGG 288
Db 121 AAGAAATCATGGCCCTACCTCCCAAGTGTCAAGCATATAACAAACCAAGTCTCTGTGG 180
Qy 289 AACAAAGATGGCATTCCTCCATGGAGTCAGATATCAGGATGGGAATCTGTGATCCAATTC 348
Db 181 AACAAAGATGGCATTCCTCCATGGAGTCAGATATCAGGATGGGAATCTGTGATCCAATTC 240
Qy 349 CCTGGTTTGTACTTCATCATATTTGCCAACTGCAGTTTCTTGTACAAATGCCAAATTCCT 408
Db 241 CCAGGTTGGTACTTCATCATCTGCCAACTGCAGTTTCTTGTGAAATGCCAGAGCATGCC 300
Qy 409 GTCGATCTCAAGTTGGAGCTTCTCATCAACAGCATATCAAAAACAGGCCCTGGTGACA 468
Db 301 GTCGACCTCAAGCTGGAGCTTCTCATCAACAAAGAGCTTAAAGAGCAGACCTGGTGACA 360
Qy 469 GTGTGTGAGTCTGGAAATGCAAAACGAAACAGATATACAGAAATCTCTCAATTTCTGCTG 528
Db 361 GTGTGTGAGTCTGGAGCACAAACCAACAGACATATACAGAACTCTCCAGTCTTGTGTG 420
Qy 529 GATTACCTCCAGGTCAACACCCACCATATCAGTCAATGTGGATACATTCAGTACATAGAT 588
Db 421 GAACACCTCCAGGTCAACACCCACCATATCAGTCAAGTGGTGAATTCAGATATGTGTGAT 480

Db 12 CAAGAAGTCATGGG-CTACCTCCAAGTGGCAAGCATCTAAACAAACCAAGTTGTCTTG 70

QY 288 GAACAAGATGCATTCCTCAATGAGTGCAGATATCAGGATGGGAATCTGGTGATCCAAAT 347

Db 71 GAACAAGATGCATTCCTCAATGAGTGCAGATATCAGGATGGGAATCTGGTGATCCAAAT 130

QY 348 CCCTGGTTGTACTTCATCATTTGGCAACTGCAGTTTCTTTGACATGCGCCAAATATTC 407

Db 131 CCCTGGTTGTACTTCATCATTTGGCAACTGCAGTTTCTTTGACATGCGCCAAATATTC 190

QY 408 TGTGATCTGAAGTTGGAGTCTCTCATCAACAGCATATCAAAAACACAGCCCTGGTGAC 467

Db 191 TGTGATCTGAAGTGGAGCTCTCATCAACAGCATATCAAAAACACAGCCCTGGTGAC 250

QY 468 AGTGTGAGTCTGGAATGCAACGAACACGATATACCAAGATCTCTCTCAATTCCTGTCT 527

Db 251 AGTGTGAGTCTGGAATGCAACGAACACGATATACCAAGATCTCTCTCAATTCCTGTCT 310

QY 528 GG 529

Db 311 GG 312

RESULT 5

AZ786450/c

LOCUS

DEFINITION 2M0032A02F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0032A02 F, DNA sequence.

ACCESSION AZ786450

VERSION AZ786450.1 GI:12924224

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 670)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

UNIVERSITY OF UTAH Genome Center

KEYWORDS Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

TELEPHONE Tel: 801 585 5606

FAX Fax: 801 585 7177

EMAIL Email: ddunn@genetics.utah.edu

INSERT LENGTH Insert Length: 10000 Std Error: 0.00

PLATE Plate: 0032 row: A column: 02

SEQ PRIMER Seq primer: CGTGTAAACGACGGCCAGT

CLASS Class: plasmid ends

High quality sequence stop: 670.

FEATURES

source

1. 670

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0032A02"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 189 a 129 c 149 g 203 t

ORIGIN

Query Match 42.4%; Score 275; DB 17; Length 670;

Best Local Similarity 80.1%; Pred. No. 2.4e-71;

Matches 323; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 246 CTCCCAAGTGGCAAGCATCTAAACAAACCAAGTTGTCTTTGGAAACAAAGATGGCATTC 305

Db 603 CTTTCCAGTGTCAAGCATCTCAACAATACCAAACTGTCTATGGAACGAAGATGGCACCAT 544

QY 306 CCATGGAGTCAGATATCAGGATGGGAATCTGGTATCCAAATCCCTGCTTGTACTTCAT 365

Db 543 CCACGGACTATATACCAAGACGGGAACTGTATAGTCCCAATTCCTGGCTTGTACTTCAT 484

QY 366 CATTTGCCAACTGCAGTTTCTTTGACAAATGCCAAATAATCTCTCGATCTGAAAGTTGGA 425

Db 483 CGTTTGCCAACTGCAGTTTCTCGTGCAGTGTCTCAATCATCTCTGGACCTGACATGCA 424

QY 426 GCTTCTCATCAACAGCATATCAAAAACAGGCGCTGGTGACAGTGTGTGAGTCTGGAAT 485

Db 423 GCTCTCATCAATTCCAAGATCAAAAACAGACAGCTGTGTAACAGTGTGTGAGTCTGGAGT 364

QY 486 GCAACGAACACAGTATACCAAGATCTCTCTCAATTCCTGCTGATACCTGCAAGTCAA 545

Db 363 TCAGAGTAAAGATCTTACCAGAAATCTCTCTCAGTTTTTGTGCTGATTAATACAGGTCAA 304

QY 546 CACCACCATATCAGTCAATGTGGATACATTCACAGTACATAGATACAAAGACCTTTCCPCT 605

Db 303 CTCTACCATATCAGTCAAGGTGGATATTTCCAGTATGTGGATACAAACACTTTCCCTCT 244

QY 606 TGAGAAATGTTGCTCCATCTTTATATACAGTAATTCAGACTGA 648

Db 243 TGATAATGCTCTATCCGCTCTCTTATATAGTAGCTCAGACTGA 201

RESULT 6

AW602310

LOCUS

DEFINITION RC5-BT0559-260100-011-G08 BT0559 Homo sapiens cDNA, mRNA sequence.

ACCESSION AW602310

VERSION AW602310.1 GI:7307049

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 276)

AUTHORS HCGP <http://www.ludwig.org.br/ORESTES>.

TITLE THE FAPESP/LICR Human Cancer Genome Project

JOURNAL Unpublished (1999)

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC56t2-RC5-BT0559->

AW602310 276 bp mRNA linear EST 23-MAR-2000

RC5-BT0559-260100-011-G08 BT0559 Homo sapiens cDNA, mRNA sequence.

AW602310

AW602310.1 GI:7307049

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 276)

HCGP <http://www.ludwig.org.br/ORESTES>.

THE FAPESP/LICR Human Cancer Genome Project

Unpublished (1999)

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC56t2-RC5-BT0559->

260100-011-G08&t3-2000-01-26&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 276.

FEATURES

source
1. .276
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0559"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 80 a 64 c 55 g 77 t

ORIGIN

Query Match 41.3%; Score 267.4; DB 10; Length 276;
Best Local Similarity 99.6%; Pred. No. 2.8e-69;
Matches 268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 312 AGTCAGATATCAGGATGGGAATCTGTGATCCCAATTCCTGGTTTCTACTTCATCATTTG 371

Db 1 AGTCAGATATCAGGATGGGAATCTGTGATCCCAATTCCTGGTTTCTACTTCATCATTTG 60

QY 372 CCNACTGCGAGTTCTTGTACATGCCCAATAATTCCTGCGATCTCAAGTTGAGCTTCT 431

Db 61 CCNACTGCGAGTTCTTGTACATGCCCAATAATTCCTGCGATCTCAAGTTGAGCTTCT 120

QY 432 CATCAACAGCATATCAAAAAACAGGCCCTGGTGACAGTGTGTGAGTCTGGGAATGCAAA 491

Db 121 CATCAACAGCATATCAAAAAACAGGCCCTGGTGACAGTGTGTGAGTCTGGGAATGCAAA 180

QY 492 GAAACACGTATACCAAGATCTCTCTCAATTCCTGGTGGATACCTGAGTCAACACCAC 551

Db 181 GAAACACGTATACCAAGATCTCTCTCAATTCCTGGTGGATACCTGAGTCAACACCAC 240

QY 552 CATATCAGTCAATGTGGATACATTCAGT 580

Db 241 CATATCAGTCAATGTGGATACATTCAGT 269

RESULT 7

BE073509
LOCUS RC5-BT0559-140200-012-E03 BT0559 Homo sapiens CDNA, mRNA sequence.
DEFINITION BE073509
ACCESSION BE073509
VERSION BE073509.1 GI:8420555
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 277)

REFERENCE

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,P.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663

TITLE

JOURNAL Laboratory of Cancer Genetics
MEDLINE Ludwig Institute for Cancer Research
COMMENT Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC5-BT0559-140200-012-E03&t3=2000-02-14&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 277.

FEATURES

source
1. .277
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT0559"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 80 a 63 c 56 g 78 t

ORIGIN

Query Match 39.3%; Score 254.8; DB 10; Length 277;
Best Local Similarity 98.9%; Pred. No. 1.7e-65;
Matches 267; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 312 AGTCAGATATCAGGATGGGAATCTGTGATCCCAATTCCTGGTTTCTACTTCATCATTTG 371

Db 1 AGTCAGATATCAGGATGGGAATCTGTGATCCCAATTCCTGGTTTCTACTTCATCATTTG 60

QY 372 C-CAACTGCGAGTTCTTGTACATGCCCAATAATTCCTGCGATCTGAAGTTGGAGCTTC 430

Db 61 CGCAACTGCGAGTTCTTGTACATGCCCAATAATTCCTGCGATCTGAAGTTGGAGCTTC 120

QY 431 TCATCAACAAGCATATCAAAAAACAGGCCCTGGTGACAGTGTGTGAGTCTGGAATGCAAA 490

Db 121 TCATCAACAAGCATATCAAAAAACAGGCCCTGGTGACAGTGTGTGAGTCTGGAATGCAAA 180

QY 491 CGAAACACGTATACCAAGATCTCTCTCAATTCCTGGTGGATACCTGAGTCAACACCA 550

Db 181 CGAAACACGTATACCAAGATCTCTCTCAATTCCTGGTGGATACCTGAGTCAACACCA 240

QY 551 CCATATCAGTCAATGTGGATACATTCAGT 580

Db 241 CCATATCAGTCAATGTGGATACATTCAGT 270

RESULT 8

BE092633/c
LOCUS mac14f02.x1 Soares mouse 3NDMS Mus musculus cDNA clone
DEFINITION IMAGE:3999867 3' similar to SW:TNF8_MOUSE P32972 CD30 LICAND ;, mRNA sequence.
ACCESSION BE092633
VERSION BE092633.1 GI:12575196
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 634)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other_ESTs:

Contact:

Email:

This clone is available royalty-free through

IMAGE Consortium

MGI:

High quality sequence stop:


```
/dev_stage= 13 weeks
/lab_host="DH10B (ampicillin resistant)"
```

the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbhH19W."

BASE COUNT 145 a 153 c 101 g 125 t

ORIGIN

Query Match 13.4%; Score 86.8; DB 9; Length 524;
Best Local Similarity 88.7%; Pred. No. 8e-15;
Matches 94; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 271 AAACCAAGTGTCTTGGGAACAAGATGTCATCTCCATGGAGTCAGATATCAGGATGGG 330
|||||
Db 1 AAACCAAGTGTCTTGGGAACAAGATGTCATCTCCATGGAGTCAGATATCAGGATGGG 60

QY 331 AATCTGGTGATCAATTCCTCGTTGTGTACTTCATCATTTGGCAAC 376
|||||
Db 61 AATCTGGTGATCAATTCCTCGTTGTGTACTTCATCATTTGGCAAC 106

RESULT 14

AZ711040

LOCUS

DEFINITION RPCI-24-132J17.TVC RPCI-24 DNA 552 bp linear GSS 24-JAN-2001

RPCI-24-132J17, DNA sequence.

ACCESSION AZ711040

VERSION GSS.

KEYWORDS AZ711040.1 GI:12443209

SOURCE house mouse.

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 552)

AUTHORS Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.

TITLE Mouse BAC End Sequences from Library RPCI-24

JOURNAL Unpublished (1999)

COMMENT Other_GSSs: RPCI-24-132J17.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC

library availability, please contact Pieter de Jong

(pdejong@mail.cho.org). Clones may be purchased from BACPAC

Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end

page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 132 row: J column: 17

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..552

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-24-132J17"

/clone_lib="RPCI-24"

/sex="Male"

/cell_type="Spleen/Brain"

/note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;

RPCI-24 Mouse BAC Library produced by Pieter de Jong. The

library was cloned in the pTARBAC1 cloning vector at the

BamHI sites using MboI partially digested male C57BL/6J

DNA."

BASE COUNT 166 a 118 c 111 g 157 t

ORIGIN

Query Match 10.8%; Score 70; DB 12; Length 463;

Best Local Similarity 69.8%; Pred. No. 8.4e-10;

Matches 110; Conservative 0; Mismatches 45; Indels 3; Gaps 1;

QY 26 CCAGCCACCTGGGGACCAGCGAGCTATTCTTTTCCACCCAGCCACTCTGGCTC 85

Db 309 CGAGGCCCTGGAGACGACGAGTGGAGCTACTTCTACCTCAGCACCCGACTGG--- 365

Query Match 13.4%; Score 86.8; DB 17; Length 552;
Best Local Similarity 82.0%; Pred. No. 8.3e-15;
Matches 100; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 527 TGGATTACCTGCGAGGTCAACACCACCATATCAGTCAATGTGGATACATTCAGTACATAG 586
|||||
Db 1 TGCATTACTTACAGGTCAACTCTACCATATCAGTCAGGGTGGTAATTTCCAGTATGGG 60

QY 587 ATACAAGCACCTTCTCTTGGAGAAATGTGTGTCCATCTCTTATACAGTAATTCAGACT 646
|||||
Db 61 ATACAACACTTCTCCCTCTTGATAATGTGCTATCCGTCCTCTTATATAGTACCTCAGACT 120

QY 647 GA 648

Db 121 GA 122

RESULT 15

BG091233

LOCUS

DEFINITION mac14f02.y1 Soares mouse 3NbMS Mus musculus cDNA clone
IMAGE:3999867 5' similar to SW:TNF8_MOUSE P32972 CD30 LIGAND ;,
mRNA sequence.

ACCESSION BG091233

VERSION BG091233.1

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 463)

JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:1497595

Seq primer: -40RP from Gibco

High quality sequence stop: 422.

Location/Qualifiers

1..463

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:3999867"

/clone_lib="Soares mouse 3NbMS"

/sex="male"

/tissue_type="Spleen"

/dev_stage="4 weeks"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5',

TGTACCAATCTCAAGTGGAGCGCGCTGTTTTTTTTTTTTTTTTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. RNA

provided by Dr. Bertrand Jordan. Library went through

three rounds of normalization, and was constructed by

Bento Soares and M.Fatima Bonaldo."

BASE COUNT 123 a 131 c 122 g 87 t

ORIGIN

Query Match 10.8%; Score 70; DB 12; Length 463;

Best Local Similarity 69.8%; Pred. No. 8.4e-10;

Matches 110; Conservative 0; Mismatches 45; Indels 3; Gaps 1;

QY 26 CCAGCCACCTGGGGACCAGCGAGCTATTCTTTTCCACCCAGCCACTCTGGCTC 85

Db 309 CGAGGCCCTGGAGACGACGAGTGGAGCTACTTCTACCTCAGCACCCGACTGG--- 365

Oy 86 TGTGCTTGTCTTCACGGTGGCCACTATTATAGGTGTGTGGTCGTTTCAGAGGACGGACTCCA 145
|||||
Db 366 TGTGCTTGTGTGGCAGTGGCGATCATTCGTGTACTGTGTAGTCCAGAAAAAGGACTCCA 425
|||||
Oy 146 TTCCCAACTCACCTGACACAGTCCCCCTCAAAGGAGGA 183
|||||
Db 426 CTCCAAATACAACGTGAAAAAGCCCTCCCTTAAAGGAGGA 463
|||||

Search completed: December 16, 2002, 07:29:53
Job time : 2228 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:21:56 ; Search time 13 Seconds
(without alignments)
1186.861 Million cell updates/sec

Title: US-09-628-126-2_COPY_19_390

Perfect score: 2080

Sequence: 1 FPQDRPFEDTCHGNPSHYD.....PVALSSTGKPVLDAGPVLFV 372

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2074	99.7	595	1	TNR8_HUMAN
2	693	33.3	498	1	TNR8_MOUSE
3	618	29.7	493	1	TNR8_RAT
4	225.5	10.8	461	1	TR1B_HUMAN
5	216	10.4	435	1	TNR3_HUMAN
6	200.5	9.6	474	1	TR1B_MOUSE
7	191.5	9.2	1587	1	LMG3_HUMAN
8	189	9.1	471	1	TR1A_BOVIN
9	184.5	8.9	415	1	TNR3_MOUSE
10	182	8.8	300	1	TR6B_HUMAN
11	176.5	8.5	461	1	TR1A_PIG
12	176	8.5	3767	1	MUA3_CAEEL
13	174.5	8.4	655	1	TR21_MOUSE
14	172.5	8.3	1696	1	PK5_BRACL
15	171.5	8.2	5376	1	ZAN_MOUSE
16	169	8.1	655	1	TR21_HUMAN
17	166.5	8.0	830	1	SREC_HUMAN
18	165	7.9	2321	1	NTC3_HUMAN
19	165	7.9	2715	1	G156_PARPR
20	164.5	7.9	351	1	CRMB_COWPX
21	164.5	7.9	1798	1	LMB2_HUMAN
22	164	7.9	325	1	V72_SFVKA
23	159	7.6	625	1	TR1L_MOUSE
24	159	7.6	2871	1	TR1L_BOVIN
25	158	7.6	401	1	TR1B_HUMAN
26	158	7.6	616	1	TR1L_HUMAN
27	158	7.6	2704	1	G168_PARPR
28	157.5	7.6	1680	1	FUR2_DROME
29	156.5	7.5	349	1	CRMB_CAMPS
30	156.5	7.5	1955	1	AGRI_CHICK
31	156	7.5	2871	1	TR1L_PIG
32	155.5	7.5	5179	1	MUC2_HUMAN
33	154.5	7.4	327	1	TNR6_MOUSE

34	154.5	7.4	2318	1	NTC3_MOUSE
35	153.5	7.4	349	1	CRMB_VARV
36	152	7.3	283	1	TR14_HUMAN
37	152	7.3	2871	1	FBN1_HUMAN
38	151.5	7.3	3718	1	LMAS_MOUSE
39	151	7.3	401	1	TR1B_RAT
40	151	7.3	1801	1	LMB2_RAT
41	150	7.2	180	1	TR22_MOUSE
42	150	7.2	326	1	VT2_MXVL
43	150	7.2	455	1	TR1A_HUMAN
44	149.5	7.2	259	1	TR10C_HUMAN
45	148.5	7.1	1799	1	LMB2_MOUSE

ALIGNMENTS

RESULT 1
TNR8_HUMAN STANDARD; PRT; 595 AA.
AC P28908;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 8 precursor (CD30L receptor) (Lymphocyte activation antigen CD30) (KI-1 antigen).
DE TNFRSF8 OR CD30.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RC TISSUE=Lymphoid;
RX MEDLINE=92154659; PubMed=1310894;
RA Duerkop H., Latza U., Hummel M., Eitelbach F., Seed B., Stein H.;
RT "Molecular cloning and expression of a new member of the nerve growth factor receptor family that is characteristic for Hodgkin's disease.";
RT Cell 68:421-427(1992).
RL [2]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RX MEDLINE=95089787; PubMed=7527901;
RA Jung W., Krueger S., Renner C., Gause A., Sahin U., Trumper L.,
RA Pfeundschuh M.;
RT "Opposite effects of the CD30 ligand are not due to CD30 mutations: results from cDNA cloning and sequence comparison of the CD30 antigen from different sources.";
RT Mol. Immunol. 31:1329-1334(1994).
RN [3]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RX MEDLINE=96437016; PubMed=8839832;
RA Horie R., Ito K., Tatewaki M., Nagai M., Aizawa S.,
RA Higashihara M., Ishida T., Inoue J., Takizawa H., Watanabe T.;
RT "A variant CD30 protein lacking extracellular and transmembrane domain is induced in HL-60 by tetracycline and is expressed in alveolar macrophages.";
RT Blood 88:2422-2432(1996).
RL [4]
CC -!- FUNCTION: Receptor for TNFRSF8/CD30L. May play a role in the regulation of cellular growth and transformation of activated lymphoblasts. Regulates gene expression through activation of NF-kappaB.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (long isoform); Cytoplasmic (short isoform).
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms: a long form (shown here) and a short form/cytoplasmic/CD30V; are produced by alternative initiation.
CC -!- PTM: PHOSPHORYLATED ON SERINE AND TYROSINE RESIDUES.
CC -!- DISEASE: MOST SPECIFIC HODGKIN'S DISEASE ASSOCIATED ANTIGEN.
CC -!- SIMILARITY: CONTAINS 6 TNFR-CYS REPEATS.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD30 entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd30.htm".


```

FT DISULFID 29 44 BY SIMILARITY.
FT DISULFID 45 58 BY SIMILARITY.
FT DISULFID 48 65 BY SIMILARITY.
FT DISULFID 69 81 BY SIMILARITY.
FT DISULFID 84 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 107 121 BY SIMILARITY.
FT DISULFID 128 146 BY SIMILARITY.
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 498 AA; 53216 MW; 98CA2A05B38AFA71 CRC64;

Query Match 33.3%; Score 693; DB 1; Length 498;
Best Local Similarity 41.4%; Pred. No. 1.3e-33;
Matches 155; Conservative 33; Mismatches 84; Indels 102; Gaps 10;

QY 1 FQDRPFEDTCGNPSHYDKAVRRCYRCPMGLFTQCPQRPDTCRKCQCEPDYVLDEA 60
DB 19 FPTDRPLKTTTCAGDLSHYPGEAARNCCYQCPGSLPTQCPGPAHCRKQCADYVYV 78

QY 61 DRCCTACVTCSDRLDVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHSHVCPAGMIVK 120
DB 79 GKCTACVTC-LPGLVEKAPCGSNPRICECPGMCCTPAVNSCARCKLH--CSGEEVVK 135

QY 121 FPGTAQNTVCEPASPGVSPACASPENCKEPPSSGTIPQAKPTVPSPATSSA-STMPVRGG 179
DB 136 SPGTAQNTVCEPASPGVSPACASPENCKEPPSSGTIPQAKPTVPSPATSSA-STMPVRGG 179

QY 180 TRLAQAAKSLTRAPDSPSSVGR-PSDFGLSPTQCPGSGDCRKCQCEPDYVLDEAGRC 238
DB 195 TNLVQEDATELVKVPESSSKAREPSPDPG----- 224

QY 239 TACVSCSRDDLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHSHVCPAGMIVK 298
DB 225 TACVSCSRDDLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHSHVCPAGMIVK 298

QY 299 MAEKDTTFEAPPLGTQDCNPTPENGAPASTSPTOSLLVDQSKTLPIPTSPAPVALSS 358
DB 225 NAEKNWTLPELPGTULPDIS-TSENSKEFASTASTLSLVDAWTSRMQ-PTS-PL---S 278

QY 359 TKGPVLADGAPVLFW 372
DB 279 TGTFFLDGPGVLFW 292

```

RESULT 3

```

ID TNR8_RAT STANDARD; PRT; 493 AA.
AC P97525;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 8 precursor (CD30L
DE receptor) (Lymphocyte activation antigen CD30).
GN TNFRSF8 OR CD30.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WKAH; TISSUE=T-cell lymphoma;
RX MEDLINE=97136705; PubMed=8982082;
RA Aizawa S., Satoh H., Horie R., Ito K., Choi S.H., Takeuchi H.,
RA Watanabe T.;
RT "Cloning and characterization of a cDNA for rat CD30 homolog and
RT chromosomal assignment of the genomic gene.";
RL Gene 182:155-162(1996).
CC -!- FUNCTION: Receptor for TNFSF8/CD30L. May play a role in
CC the regulation of cellular growth and transformation of
CC activated lymphoblasts. Regulates gene expression through
CC activation of NF-kappaB (By similarity).

```

```

CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Very low level of expression. Detected in
CC spleen, thymus and lung. Highly expressed in HTLV-1 infected T-
CC cell lines.
CC -!- INDUCTION: By phytohemagglutinin (PHA) in spleen T-cells.
CC -!- SIMILARITY: CONTAINS 3 TNFR-CYS REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: D42117; BAA07699.1; -.
DB InterPro: IPR001368; TNFR_c6.
DB Pfam: PF00020; TNFR_c6; 3.
DB SMART: SM00208; TNFR; 2.
DB PROSITE: PS00652; TNFR_NGFR_1; FALSE_NEG.
DB PROSITE: PS00030; TNFR_NGFR_2; 1.
KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 493 TUMOR NECROSIS FACTOR RECEPTOR
FT DOMAIN 19 255 SUPERFAMILY MEMBER 8.
FT DOMAIN 256 276 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 277 493 POTENTIAL.
FT REPEAT 28 66 CYTOPLASMIC (POTENTIAL).
FT REPEAT 68 106 TNFR-CYS 1.
FT REPEAT 107 150 TNFR-CYS 2.
FT REPEAT 107 150 TNFR-CYS 3.
FT DISULFID 29 44 BY SIMILARITY.
FT DISULFID 45 58 BY SIMILARITY.
FT DISULFID 84 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 107 121 BY SIMILARITY.
FT DISULFID 128 142 POTENTIAL.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 493 AA; 52658 MW; 40BE75033B3CBB7C CRC64;

Query Match 29.7%; Score 618; DB 1; Length 493;
Best Local Similarity 38.1%; Pred. No. 2.7e-29;
Matches 142; Conservative 35; Mismatches 92; Indels 104; Gaps 8;

QY 1 FQDRPFEDTCGNPSHYDKAVRRCYRCPMGLFTQCPQRPDTCRKCQCEPDYVLDEA 60
DB 19 FPKDRPLDTCGDLSTYYPGEAARNCCYQCPGSLPTQCPGPAHCRKQCADYVYV 78

QY 61 DRCCTACVTCSDRLDVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHSHVCPAGMIVK 120
DB 79 GKCTACVTC-LPGLVEKAPCGSNPRICECPGMCCTPAVNSCARCKLH--CSGEEVVK 131

QY 121 FPGTAQNTVCEPASPGVSPACASPENCKEPPSSGTIPQAKPTVPSPATSSA-STMPVRGG 180
DB 132 FPDTAENATICEPLSPGSGNGSNPDDCKTLTSHHTTPOAITPLETSPANDSVKSLLPKQVT 191

QY 181 RLQAQAAKSLTRAPDSPSS-VGRPSDDPLSPTQCPGSGDCRKCQCEPDYVLDEAGRC 239
DB 192 DFVNEGATKLKVPESSSKAREPSPDPG----- 220

QY 240 ACVSCSRDDLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHSHVCPAGMIVK 299
DB 221 ACVSCSRDDLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHSHVCPAGMIVK 299

QY 300 AEKDTTFEAPPLGTQDCNPTPENGAPASTSPTOSLLVDQSKTLPIPTSPAPVALSS 359
DB 222 AEMNTLKLPPPGVTPDIS-TSENSMEFASTASTLSLVDAWTSRMQ-PTS-PL---ST 275

QY 360 GKPVLDAGPVLFW 372
DB 276 GTFPLDGTMLFW 288

```

RESULT 4

ID TRIB_HUMAN STANDARD: PRT: 461 AA.

AC P20333; Q16042;

DT 01-FEB-1991 (Rel. 17, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Tumor necrosis factor receptor superfamily member 1B precursor (Tumor necrosis factor receptor 2) (p80) (TNF-R2) (p75) (CD120b) (Etanercept)

DE [Contains: Tumor necrosis factor binding protein 2 (TBPII)].

CN TNFRSF1B OR TNFR2 OR TNFR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

[1]

RN SEQUENCE FROM N.A.

RP MEDLINE=90260639; PubMed=2160731;

RA Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R., Dower S.K., Cosman D., Goodwin R.G.;

RT "A receptor for tumor necrosis factor defines an unusual family of cellular and viral proteins.";

RL Science 248:1019-1023(1990).

[2]

RN SEQUENCE FROM N.A.

RP MEDLINE=91045991; PubMed=2172983;

RA Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W., Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;

RT "A second tumor necrosis factor receptor gene product can shed a naturally occurring tumor necrosis factor inhibitor.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).

[3]

RN SEQUENCE FROM N.A.

RP MEDLINE=9629745; PubMed=8661109;

RA Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J., Lepassier D., Stallard B.J., Goeddel D.V., Desauvage F.J., Brodeur G.M.;

RT "Physical mapping and genomic structure of the human TNFR2 gene.";

RL Genomics 35:94-100(1996).

[4]

RN SEQUENCE OF 37-461 FROM N.A.

RP MEDLINE=91370690; PubMed=1966549;

RA Dembic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R., Brockhaus M., Lesslauer W.;

RT "Two human TNF receptors have similar extracellular, but distinct intracellular, domain sequences.";

RL Cytokine 2:231-237(1990).

[5]

RN SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.

RP MEDLINE=90349572; PubMed=2166946;

RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D., Ringold G.M.;

RT "Complementary DNA cloning of a receptor for tumor necrosis factor and demonstration of a shed form of the receptor.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).

[6]

RN SEQUENCE OF 27-31.

RP MEDLINE=9010215; PubMed=2153136;

RA Engelmann H., Novick D., Wallach D.;

RT "Two tumor necrosis factor-binding proteins purified from human urine. Evidence for immunological cross-reactivity with cell surface tumor necrosis factor receptors.";

RL J. Biol. Chem. 265:1531-1536(1990).

[7]

RN SEQUENCE OF 23-40; 65-69; 136-141; 300-306 AND 346-362.

RP MEDLINE=91056048; PubMed=2173696;

RA Loetscher H., Schlaeger E.J., Lahm H.W., Pan Y.-C.E., Lesslauer W., Brockhaus M.;

RT "Purification and partial amino acid sequence analysis of two distinct tumor necrosis factor receptors from HL60 cells.";

RL J. Biol. Chem. 265:20131-20138(1990).

[8]

RN CHARACTERIZATION.

RP MEDLINE=93016040; PubMed=1328224;

RA Pennica D., Lam V.T., Mize N.K., Weber R.F., Lewis M., Fendly B.M., Lipari M.T., Goeddel D.V.;

RT "Biochemical properties of the 75-kDa tumor necrosis factor receptor. Characterization of ligand binding, internalization, and receptor phosphorylation.";

RL J. Biol. Chem. 267:21172-21178(1992).

[9]

RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN COMPLEX WITH TNF2.

RP MEDLINE=99221490; PubMed=10206649;

RA Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;

RT "Structural basis for self-association and receptor recognition of human TNF2.";

RL Nature 398:533-538(1999).

CC -!- FUNCTION: Receptor with high affinity for TNFSF2/TNF-alpha and approximately 5-fold lower affinity for homotrimeric TNFSF1/lymphotoxin-alpha.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein and secreted.

CC -!- PTM: Phosphorylated; mainly on serine residues and with a very low level on threonine residues.

CC -!- PTM: A soluble form (tumor necrosis factor binding protein 2) is produced from the membrane form by proteolytic processing.

CC -!- PHARMACEUTICAL: Available under the name Enbrel (Immunex and Wyeth-Ayerst). Used to treat moderate to severe rheumatoid arthritis (RA). Enbrel consist of the extracellular ligand-binding portion of TNFR2 linked to an immunoglobulin Fc chain. It binds to TNF-alpha and blocks its interactions with receptors.

CC -!- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.

CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD120b entry; WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd120b.htm".

CC -!- DATABASE: NAME=Enbrel; NOTE=Clinical information on Enbrel; WWW="http://www.enbrelinfo.com/".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M32315; AAA59929.1; -.

DR EMBL; U52165; AAC50622.1; -.

DR EMBL; U52156; AAC50622.1; JOINED.

DR EMBL; U52157; AAC50622.1; JOINED.

DR EMBL; U52158; AAC50622.1; JOINED.

DR EMBL; U52159; AAC50622.1; JOINED.

DR EMBL; U52160; AAC50622.1; JOINED.

DR EMBL; U52161; AAC50622.1; JOINED.

DR EMBL; U52162; AAC50622.1; JOINED.

DR EMBL; U52163; AAC50622.1; JOINED.

DR EMBL; U52164; AAC50622.1; JOINED.

DR EMBL; M55994; AAA36755.1; -.

DR EMBL; S63368; AAB19824.2; -.

DR EMBL; M35857; AAB63262.1; -.

DR PIR; A35356; A35356.

DR PIR; A36007; A36007.

DR PIR; A36475; A36475.

DR PIR; B35010; B35010.

DR PIR; A23666; A23666.

DR PDB; 1CA9; 12-APR-99.

DR Genew; HGNC:11917; TNFRSF1B.

DR MIM; 191191; -.

DR InterPro; IPR001388; TNFR_c6.

DR Pfam; PF00020; TNFR_c6; 4.

DR ProDom; PD000771; TNFR_c6; 1.

DR SMART; SM00208; TNFR; 4.

DR PROSITE; PS00652; TNFR_NGFR_1; 2.

DR PROSITE; PS00500; TNFR_NGFR_2; 4.

DR Receptor; Transmembrane; Glycoprotein; Repeat; Signal;

KW

```

KW Phosphorylation; Pharmaceutical; 3D-structure.
FT SIGNAL 1 22
FT CHAIN 23 461
FT TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER 1B, MEMBRANE FORM.
FT CHAIN 27 257
FT DOMAIN 23 287
FT TRANSMEM 258 461
FT DOMAIN 288 461
FT REPEAT 39 76
FT REPEAT 77 118
FT REPEAT 119 162
FT REPEAT 163 201
FT REPEAT 201 461
FT DISULFID 40 53
FT DISULFID 54 67
FT DISULFID 57 75
FT DISULFID 78 93
FT DISULFID 96 110
FT DISULFID 100 118
FT DISULFID 120 126
FT DISULFID 134 143
FT DISULFID 137 161
FT DISULFID 164 179
FT CARBOHYD 171 171
FT CARBOHYD 193 193
FT CONFLICT 141 141
FT CONFLICT 196 196
FT CONFLICT 363 363
FT CONFLICT 461 461
FT SEQUENCE 461 AA; 48291 MW; 603D0AEICD69ACBF CRC64;

Query Match 10.8%; Score 225.5; DB 1; Length 461;
Best Local Similarity 28.0%; Pred. No. 1.1e-06;
Matches 61; Conservative 24; Mismatches 114; Indels 19; Gaps 6;

QY 18 YDKAVRCCYRCMGLFTPOCQORPTDCRKOCEPDYL---DEADRTACTVTCRRDDL 74
D 111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 YDQTAQMCSCSPQGHAKVCTKTSDFVCDSCSDSTYTLQWNVPECLSCGRCSSDQ 104
QY 75 VEKTPCAWNSRVCRCPCMGFCSTSAVNSCARCFHFHVCAGVIFKPGTAQKNTVCEPA 134
D 111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 VETQACTREQNICRCPGWYCALSKQEGCRGLCAPLRKCRPGFVARPGTSDVYCKPC 164
QY 135 SPGV-SPACASPENCKEPCSGTIPQAKPTVPSPATSSASTMPVR---GTRLAQEAASK 189
D 111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 165 AFGTFSNTSSDIDCR-PHQICNVVAIPGNASMDVACTSTSPTRSMAPGAVHLPPQVSTR 223
QY 190 LTRADPSVSGRPSDPLGSL-----PTQCPGEGSD 221
D 111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 224 SQHTQTP-----EPSTAPSTSFLLPMGPSPPAEGSTGD 257

RESULT 5
TNR3_HUMAN
ID TNR3_HUMAN STANDARD; PRT; 435 AA.
AC P36941.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 3 precursor
DE (Lymphotoxin-beta receptor) (Tumor necrosis factor receptor 2 related
DE protein) (Tumor necrosis factor C receptor).
GN LTBR OR TNFRSF3 OR TNFCR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93252381; PubMed=8486360;
RA Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;
RT "Construction and evaluation of a hncDNA library of human 12p
RT transcribed sequences derived from a somatic cell hybrid.";
RG Genomics 16:214-218(1993).

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP FUNCTION.
RX MEDLINE=94225209; PubMed=8171323;
RA Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,
RA Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;
RT "A lymphotoxin-beta-specific receptor.";
RL Science 264:707-710(1994).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=9923511; PubMed=10207006;
RA Wu M.-Y., Wang P.-Y., Han S.-H., Hsieh S.-L.;
RT "The cytoplasmic domain of the lymphotoxin-beta receptor mediates cell
RT death in HeLa cells.";
RL J. Biol. Chem. 274:11868-11873(1999).
RN [5]
RP FUNCTION.
RX MEDLINE=20261554; PubMed=10799510;
RA Rooney I.A., Butrovich K.D., Glass A.A., Borboroglu S., Benedict C.A.,
RA Whitebeck J.C., Cohen G.H., Eisenberg R.J., Ware C.F.;
RT "The lymphotoxin-beta receptor is necessary and sufficient for
RT LIGHT-mediated apoptosis of tumor cells.";
RL J. Biol. Chem. 275:14307-14315(2000).
CC -1- FUNCTION: Receptor for the heterotrimeric lymphotoxin containing
CC LTA and LTB, and for TNFSF14/LIGHT. Promotes apoptosis via TRAF3
CC and TRAF5. May play a role in the development of lymphoid organs.
CC -1- SUBUNIT: Self-associates.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L04270; AAA36757.1; -.
DR EMBL; BC026262; AAH26262.1; -.
DR HSSP; P25942; ICDF.
DR Genes; HGNC:6718; LTBR.
DR MIM; 600979; -.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 4.
DR ProDom; PD000771; TNFR_c6; 1.
DR SMART; SM00208; TNFR; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
DR Receptor; Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 30
FT CHAIN 31 435
FT TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER 3.
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 228 248
FT DOMAIN 249 435
FT REPEAT 42 81
FT REPEAT 82 124
FT REPEAT 125 168
FT REPEAT 169 211
FT DISULFID 43 58
FT DISULFID 59 72
FT DISULFID 62 80
FT DISULFID 83 98
FT DISULFID 101 116
FT DISULFID 104 124
FT DISULFID 126 132
FT DISULFID 139 148
FT DISULFID 142 167

```



```

FT DISULFID 170 185 BY SIMILARITY.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 435 AA; 46709 MW; 624626E6022F656F CRC64;

Query Match 10.4%; Score 216; DB 1; Length 435;
Best Local Similarity 23.2%; Pred. No. 3.7e-06;
Matches 82; Conservative 30; Mismatches 112; Indels 130; Gaps 14;

Qy 10 TCHGNPSHYDKAVRCRCYRCMPGLFPTQCPQ-RPTDCRKOCEPDYILDEADRCTACTVT 68
Db 42 TCRDQKEYEYEQHRCICRCPGPGYVSAKCSIRDVTC-ATCAENSNEHWNITICOL 100
Qy 69 CSRDP-----LVEKTEPCAWNSSRVCECRPGMFCSTSAV
Db 101 CRPCDPVMGLEIEIAPCTSKRKTQCRQCPGMFCAAWALECTHCELLSDCPGPEALKDREV 160
Qy 102 ----NSC-----ARCFHVSVC-AGMIVKFCPTAQKNTVC-----EPASPG 137
Db 161 GKGNNHCVPCKAGHFQNTSSPSARCOPTRCENQGLVEAAPGTAQSDTCKKNPLEPLPPE 220
Qy 138 VSP-----AC-----ASPENCKE-----PSSGTIPQ 158
Db 221 MSGTMLAVLLPTAFFLLLATVFCIWKSHSCLKGLSLKRRPQEGPNPVAGSWEF 280
Qy 159 AKTP-----VSPATSSASTMPVGGTRLAQAAASKLTRAPD-----SPSS 199
Db 281 PKAHYFPDLVQPLPLISGDSVPSTGLPAADVLEAGVPQQSPDLTRPQLPEGEQSQ 340
Qy 200 VGRPSSDGL-----SPTQCPGSGDCKRCQCPDYILDEAG 236
Db 341 VAHGTN--GIHVTGSGMTITGNIYINGPVLGGPGPGDLPATPEPPYPIPEG 392

RESULT 6
TRIM_MOUSE STANDARD; PRT; 474 AA.
ID TRIM_MOUSE
AC P25119; P97893;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 1B precursor (Tumor
DE necrosis factor receptor 2) (TNF-R2) (p75).
GN TNFRSF1B OR TNFR2 OR TNFR-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9118785; PubMed=1849278;
RA Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,
RA Wong G.H., Chen E.Y., Goeddel D.V.;
RT "Cloning and expression of cDNAs for two distinct murine tumor
RT necrosis factor receptors demonstrate one receptor is species
RT specific."
RL Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91246168; PubMed=1645445;
RA Goodwin R.G., Anderson D., Jerzy R., Davis T., Brannan C.I.,
RA Copeland N.G., Jenkins N.A., Smith C.A.;
RT "Molecular cloning and expression of the type 1 and type 2 murine
RT receptors for tumor necrosis factor."
RL Mol. Cell. Biol. 11:3020-3026 (1991).
RN [3]
RP SEQUENCE OF 1-26 FROM N.A.
RC STRAIN=NOD;
RA Jacob C.O., Liu J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-22 FROM N.A.
RC TISSUE=Liver;

```

```

RA Kissergerhis M., Fellowes R., Feldmann M., Chernajovsky Y.;
RL FUNCTION (MAY-1995) to the EMBL/GenBank/DBJ databases.
-!- SUBMITTER: Receptor with high affinity for TNFSF2/TNF-alpha and
CC approximately 5-fold lower affinity for homotrimeric
CC TNFSF1/lymphotoxin-alpha (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M60469; AAA39752.1; -
DR EMBL: M59378; AAA0463.1; -
DR EMBL: U39488; AAA8502.1; -
DR EMBL: X87128; AAA60618.1; -
DR PIR: B38634; B38634.
DR HSP: P19438; INCF.
DR MGD: MGI:1314883; Tnfrsf1b.
DR InterPro: IPR001368; TNFR_C6.
DR Pfam: PF00020; TNFR_C6; 4.
DR ProDom: PD000771; TNFR_C6; 1.
DR SMART: SM00208; TNFR; 4.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 3.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 474 TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER 1B.
FT DOMAIN 23 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 259 288 POTENTIAL.
FT DOMAIN 289 474 CYTOPLASMIC (POTENTIAL).
FT REPEAT 39 77 TNFR-CYS 1.
FT REPEAT 78 119 TNFR-CYS 2.
FT REPEAT 120 164 TNFR-CYS 3.
FT REPEAT 165 203 TNFR-CYS 4.
FT DISULFID 40 54 BY SIMILARITY.
FT DISULFID 55 68 BY SIMILARITY.
FT DISULFID 58 76 BY SIMILARITY.
FT DISULFID 79 94 BY SIMILARITY.
FT DISULFID 97 111 BY SIMILARITY.
FT DISULFID 101 119 BY SIMILARITY.
FT DISULFID 121 127 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 139 163 BY SIMILARITY.
FT DISULFID 166 181 BY SIMILARITY.
FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 474 AA; 50319 MW; 462EAE398C4D6563 CRC64;

Query Match 9.6%; Score 200.5; DB 1; Length 474;
Best Local Similarity 20.5%; Pred. No. 3.2e-05;
Matches 93; Conservative 39; Mismatches 160; Indels 161; Gaps 16;

Qy 18 YDKAVRCRCYRCMPGLFPTQCPQRPPTDCRKOCEPDYILDEADRCTACTVCRD---DL 74
Db 46 YDRKAQMCCKAKPGPGYVVKHFCNKSTDTVCADCEASMTYQVWNQFRTCLSCSSCTTDQ 105
Qy 75 VEKTPCAWNSRVCECRPGMFCSTSA-VNSCARCFHVSVC-PAGM----- 117
Db 106 VEIRACTKQONRVCAACEAGRYCALKTHSGSCRCQCMRLSKCGPGFGVASSRAPNGVLCKA 165
Qy 118 -----IVKPPGTAQKNTVCEPASPGVSPA-----CASPENCK 149
Db 166 CAPGTFSDTTSSTDVCRPHRICSLAIPGNASTDVCAPESPTLSAIPRLYVSQEPTR 225
Qy 150 EPSSTGTPQAKPTVPSPATSSAST-----MPVRGTRL----- 182
Db 226 SQPLDQEPGPSQTP-SILTSLGSTPIIEQSTKGGISLPIGLIVGVTSLGLMLGLVNCII 284

```

QY 183 -----AQEAAS-----KLTRAPDSSPVGRPSSDPGLSPTQ 213
 Db 285 LVORKKKPSCLORDAKVPHVDEKSDQAVGLEQOHLITAPSSSSSSLESSASAG--DRR 342
 QY 214 PCPEGSGDCRKOCEPDYIIDEAGRCTACVSCSRDDILVKTPCAWNSRICECRPGM----- 269
 Db 343 APGGHQARVMAEAGQFOE-----ARASRISDSSHGSHGTH 380
 QY 270 ICATSATNSCARCVPIPCAGETVTRPQDMAEKDITFEAPLGTQPCDNP--TPENGCEAP 327
 Db 381 VNVTCIVNVCSSDSSHQSSQASATVGD-----PDAKPSAPKDEQVP 424
 QY 328 ASTS--PTQSLVLVDSQASKTL-----PIPTSAP 353
 Db 425 FSQEECPSSQ--PCETTETLQSHERPLPLGVP 454

RESULT 7
 LMG3_HUMAN
 ID LMG3_HUMAN STANDARD; PRT; 1587 AA.
 AC Q9Y6N6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin gamma-3 chain precursor (Laminin 12 gamma 3).
 GN LAMC3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=99242614; PubMed=10225960;
 RA Koch M., Olson P.F., Albus A., Jin W., Hunter D.D., Brunken W.J.,
 RA Burgesson R.E., Champlaud M.F.;
 RT "Characterization and expression of the laminin gamma3 chain: a novel,
 RL J. Cell Biol. 143:603-618(1999).
 CC -!- FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration, and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC -!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end.
 CC THE GAMMA-3 CHAIN IS A SUBUNIT OF LAMININ-12.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: Broadly expressed in: skin, heart, lung, and
 CC the reproductive tracts.
 CC -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -!- DOMAIN: DOMAIN IV IS GLOBULAR.
 CC -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -!- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF041835; AAD36991.1; -;
 CC HSSP; P02468; I1LE.
 CC Genew; HGNC:6494; LAMC3.
 CC MIM; 604349; -;
 CC InterPro; IPR000561; EGF-like.

DR InterPro; IPR001886; LamNT.
 DR InterPro; IPR000034; Laminin_B.
 DR InterPro; IPR002049; Laminin_EGF.
 DR Pfam; PF000052; laminin_B; 1.
 DR Pfam; PF000053; laminin_EGF; 10.
 DR Pfam; PF000055; laminin_Nterm; 1.
 DR PRINTS; PR000111; EGF_LAMININ.
 DR PRODOM; PD002082; LamNT; 1.
 DR PRODOM; PD003031; Laminin_B; 1.
 DR SMART; SM00180; EGF_Lam; 10.
 DR SMART; SM00001; EGF_Like; 1.
 DR SMART; SM00281; LamB; 1.
 DR SMART; SM00136; LamNT; 1.
 DR PROSITE; PS00022; EGF_1; 7.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 10.
 DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 1587 LAMININ GAMMA-3 CHAIN.
 FT DOMAIN 20 270 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 271 326 LAMININ EGF-LIKE 1.
 FT DOMAIN 327 382 LAMININ EGF-LIKE 2.
 FT DOMAIN 383 429 LAMININ EGF-LIKE 3.
 FT DOMAIN 430 479 LAMININ EGF-LIKE 4.
 FT DOMAIN 480 489 LAMININ EGF-LIKE 5 (N-TERMINAL).
 FT DOMAIN 490 672 LAMININ DOMAIN IV.
 FT DOMAIN 673 706 LAMININ EGF-LIKE 5 (C-TERMINAL).
 FT DOMAIN 707 754 LAMININ EGF-LIKE 6.
 FT DOMAIN 755 809 LAMININ EGF-LIKE 7.
 FT DOMAIN 810 865 LAMININ EGF-LIKE 8.
 FT DOMAIN 866 916 LAMININ EGF-LIKE 9.
 FT DOMAIN 917 964 LAMININ EGF-LIKE 10.
 FT DOMAIN 965 1013 LAMININ EGF-LIKE 11.
 FT DOMAIN 1014 1587 DOMAIN II AND I.
 FT DOMAIN 1071 1141 COILED COIL (POTENTIAL).
 FT DOMAIN 1200 1229 COILED COIL (POTENTIAL).
 FT DOMAIN 1424 1504 COILED COIL (POTENTIAL).
 FT DOMAIN 1535 1579 COILED COIL (POTENTIAL).
 FT SITE 1059 1061 CELL ATTACHMENT SITE (POTENTIAL).
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 837 837 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1185 1185 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1518 1518 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1587 AA; 172051 MW; 3CB6E09B5F203319 CRC64;

Query Match 9.2%; Score 191.5; DB 1; Length 1587;
 Best Local Similarity 21.2%; Pred. No. 0.0003;
 Matches 99; Conservative 44; Mismatches 152; Indels 173; Gaps 24;
 QY 7 FEDTCHGNPSHY-YDKAVRCCYRCPMGLF---PTQCPQRPRTDCRK----- 49
 Db 343 FRSTGGGRCHCRDHTAGPHCERCQENFYHWDPMPC--QPCDCQSAGSLHLCDDTGT 400
 QY 50 -QCEPDYVLEADRCT-----ACVTC-----SRDDLVEKTPCAWN-SSRVCE 89
 Db 401 CACKPTVTGWKCDRLCGFHSLEGGCRPCTCNPAGSLDTCPRSRCPCCKENVEGNLCD 460
 QY 90 -CRGMEF-----CSTSAVNSCARCFHFS--VC----- 113
 Db 461 RCRPGTENLQHPNAGCSSCF-CYGHSKVKCASTAQFVHHILSDPHQGAEGWARSVGGS 519
 QY 114 -----PAGMIV-----KPGTAQKNTVCEP-----ASGVSP----- 140
 Db 520 EHSQWSPNGVLLSPEDEELTAPGKFLGD-QRESYGQPLILTRFVPPGDSPLVOLRLUE 578
 QY 141 -----ACASPENCKEPPSGSTIPQAKPTVPSPATSSASTMPVRGGTRLAQEAASK 189

```

Db 579 GTGLALSRHSSLSGPDARASGG---RAQVPLQETSEDVAPPLPPFHFQRLLANLTSL 635
Qy 190 LTRAPDPSVSG-----RPSDDPGLSPQ-----PCPEG-SGDKRQCEPDY--YL 232
Db 636 RLRSVPGSPAGVFLTEVRLTSARPLSPASWVEICSCPTGYTGQFCESCAPGVKREM 695
Qy 233 DEAGRCTACVSRDLDVEKTPCAWNSRFTCECPGMCATSATN--SCARCVPPICAG 290
Db 696 POGGPYASCV-----PCTNQHGTCDPNTGICVCSHHTGTPSGERCLP----- 738
Qy 291 ETVTRKPDMAEKDTTFEAPPLGTQDPCNTPENGEPASTSPQTQSLV 338
Db 739 -----GFYGNPFAGQADDCPCPCPGSACTIPESGEV 773

RESULT 8
TRIA_BOVIN
ID TRIA_BOVIN STANDARD; PRT; 471 AA.
AC O19131;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 1A precursor (p60)
DE (TNF-R1) (TNF-R1) (p55).
GN TNFSF1A OR TNFR1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RX MEDLINE=98273505; PubMed=9613449;
RA Lee E.-K., Kehrl M.E. Jr., Taylor M.J.;
RT "Cloning and sequencing of cDNA encoding bovine tumor necrosis factor
RT (TNF)-receptor 1.";
RL Vet. Immunol. Immunopathol. 61:379-385(1998).
CC -!- FUNCTION: Receptor for TNFSF2/TNF-alpha and homotrimeric
CC TNFSF1/lymphotoxin-alpha. The adaptor molecule FADD recruits
CC caspase-8 to the activated receptor. The resulting death-inducing
CC signaling complex (DISC) performs caspase-8 proteolytic activation
CC which initiates the subsequent cascade of caspases (aspartate-
CC specific cysteine proteases) mediating apoptosis (By similarity).
CC -!- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC PROTEINS SUCH AS TRAFs, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC NF-KAPPA B SIGNALING (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U90937; AB65143.1; -
CC HSSP; P19438; 1TNR.
CC InterPro; IPR000488; Death.
CC InterPro; IPR001368; TNFR_c6.
CC Pfam; PF00020; TNFR_c6; 3.
CC Pfam; PF00531; Death; 1.
CC ProDom; PD000771; TNFR_c6; 1.
CC SMART; SM00005; DEATH; 1.
CC SMART; SM00208; TNFR; 3.

```

```

DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS0050; TNFR_NGFR_2; 3.
DR PROSITE; PS0017; DEATH_DOMAIN; 1.
KW Receptor; Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 471 TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER 1A.
FT DOMAIN 22 210 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 211 233 POTENTIAL.
FT DOMAIN 234 471 CYTOPLASMIC (POTENTIAL).
FT REPEAT 43 82 TNFR-CYS 1.
FT REPEAT 83 125 TNFR-CYS 2.
FT REPEAT 126 166 TNFR-CYS 3.
FT REPEAT 167 195 TNFR-CYS 4.
FT DOMAIN 340 360 N-SMASE ACTIVATION DOMAIN (NSD).
FT DEATH 372 457 DEATH.
FT DISULFID 44 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 81 BY SIMILARITY.
FT DISULFID 84 99 BY SIMILARITY.
FT DISULFID 102 117 BY SIMILARITY.
FT DISULFID 105 125 BY SIMILARITY.
FT DISULFID 127 143 BY SIMILARITY.
FT DISULFID 146 166 BY SIMILARITY.
FT DISULFID 168 179 BY SIMILARITY.
FT DISULFID 182 194 BY SIMILARITY.
FT DISULFID 185 190 BY SIMILARITY.
FT CARBOHYD 54 54 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 471 AA; 51367 MW; 5243EF514DFE81C4 CRC64;

Query Match 9.1%; Score 189; DB 1; Length 471;
Best Local Similarity 24.1%; Pred. No. 0.00014;
Matches 98; Conservative 43; Mismatches 165; Indels 100; Gaps 23;

Qy 8 EDTC-HGNPSHYDKAVRRCYRCYPMGLFPTQC--PQPTDCRKCQEPDYLDADRCT 64
Db 41 ESPCPGKYNHNPONSTI--CCTCKHKGTLYNDCPGGRDTCR-VCAPGTYYTALENHLR 97
Qy 65 ACVTCs--RDDL--VKPTCAWNSRVCRCRGMFCSTSAVNSCARCFFHSCVAGMIVK 120
Db 98 RLCSRCRDMEFVETSPCVDRDTCVCGCRKNQYREYWG-ETGFRCLNCSLCPNG-TVN 155
Qy 121 FCTAOKNTVCEPASGV---SPACASPENCKEPPSSCTIPQAKPTVPSPATSSASTMPVR 177
Db 156 IPCQERQDITICH-CHMGFFLKGAKICSDCKNKEC-----EKLCPTRPST----- 200
Qy 178 GGTRLAQAASAKLTRAPDPSVSRPSSDPGLSPTQPCPEGSGDC-----RKQ 225
Db 201 -----GKDSQDPGTTVLLPLVIVFGLCLASFASVVLACRYQR 237
Qy 226 CEPDYLDDEAGRTACVSRDLDVEKTPCAWNSRFTCECPGMCATSATNSCARCV-P 284
Db 238 WKPKLYSIICQGST-LVKEGEPELLVPAP-GFNPTTT-----ICFSSTPSSPVSIIP 288
Qy 285 YPIC-----AGETVTRKPDMAEKDTTFEAPPLGTQDPCNTPENGEPAST----- 330
Db 289 YISCDRSNFCVAVSPSSSETAPPHL-KAGPILPGPPASTHL-CTPGP-----PASTHLCTP 341
Qy 331 SPTQSLVDSQASKTLPIPTSAFVALS---STGKPVLDAGPVLEW 372
Db 342 GPPASTHLCTPVOKWEASAPSPADQLADADPATLYAVVDGVPPSRW 387

RESULT 9
TNFR3_MOUSE
ID TNFR3_MOUSE STANDARD; PRT; 415 AA.
AC P50284;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

DE Tumor necrosis factor receptor superfamily member 3 precursor
GN LTBR OR TNFRSF3 OR TNFCR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CVB; TISSUE=Lung;
RX MEDLINE=96072804; PubMed=7594541;
RA Force W.R., Walter B.N., Hession C., Tizard R., Kozak C.A.,
RA Browning J.L., Ware C.F.;
RT "Mouse lymphotoxin-beta receptor. Molecular genetics, ligand binding,
RT and expression.";
RL J. Immunol. 155:5280-5288(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96163885; PubMed=8586432;
RA Nakamura T., Tashiro K., Nazarea M., Nakano T., Sasayama S.,
RA Honjo T.;
RT "The murine lymphotoxin-beta receptor cDNA: isolation by the signal
RT sequence trap and chromosomal mapping.";
RL Genomics 30:312-319(1995).
CC -!- FUNCTION: Receptor for the heterotrimeric lymphotoxin containing
CC LTA and LTB, and for TNFSF14/LIGHT. Promotes apoptosis via TRAF3
CC and TRAF5. May play a role in the development of lymphoid organs
CC (By similarity).
CC -!- SUBUNIT: Self-associates (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U29173; AAA68964.1; -;
DR EMBL; L38423; AAB00846.1; -;
DR EMBL; U30798; AAA81334.1; -;
DR HSSP; O14763; 1D0G.
DR MGD; MGI:104875; Ltbr.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 3.
DR ProDom; PD000771; TNFR_c6; 1.
DR SMART; SM00208; TNFR; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00650; TNFR_NGFR_2; 3.
KW Receptor; Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 30
FT CHAIN 31 415
FT TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER 3.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT TRANSMEM 224 244
FT DOMAIN 245 415
FT REPEAT 42 81
FT REPEAT 82 124
FT REPEAT 125 170
FT REPEAT 171 213
FT REPEAT 214 257
FT DISULFID 43 58
FT DISULFID 59 72
FT DISULFID 62 80
FT DISULFID 83 98
FT DISULFID 101 116
FT DISULFID 104 124
FT DISULFID 126 132
FT DISULFID 139 150
FT DISULFID 142 169
FT DISULFID 172 187
FT CARBOHYD 40

FT CARBOHYD 179 179 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 415 AA; 44956 MW; 29B326A566AEF661 CRC64;
Query Match 8.9%; Score 184.5; DB 1; Length 415;
Best Local Similarity 23.9%; Pred No. 0.00024;
Matches 68; Conservative 25; Mismatches 103; Indels 89; Gaps 12;
QY 2 PDORPFEDTCHGNPESHYDKAVRCCVRCVPMGLFPPTQCCPQRPTDCRQKQCEPDYVLEAD 61
DB 34 PPVRIENQTCWDODKEYEPEMHVDVCCSRCPPEFVAVCSRSQDVTCKTCHPHNSYNEHN 93
QY 62 RCTACVTCSDLLV----EKTPCAMNSSRCRCRPMFCSTSAVNSC----- 104
DB 94 HLTCTCLQPCDIVLGFEEVAPCTSDRKACRCQPGMSC-VYLDNECVHCEERLVLCP 152
QY 105 -----ARCFHVSVCV-AGMIVKPFPGTAQKNTVC 131
DB 153 GTEAEVTDIMTDVNCVCPKPGHFQNTSSPRACQPHTRCEIQGLVEAAGTYSYDTIC 212
QY 132 E-PASPGVS-----PACA---SPENCKEPSSGTI---POAKPTPV 165
DB 213 KNPEFGAMLLAILLSLVLFLLFTTVLACAWRHPSLCRK--IGTLKRHPEGESEPPC 270
QY 166 PATSSASTMPVRGGTRLAQEAASKLTRAPD-SPSSVGRPSDDPL 209
DB 271 PAP-----RADPHFDLAEPDLPMSGDLSPSPAGPPTA-PSL 306
RESULT 10
TR6B_HUMAN
ID TR6B_HUMAN STANDARD; PRT; 300 AA.
AC O95407;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 6B precursor (Decoy
GN TNFRSF6B OR DCR3 OR TR6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=99087326; PubMed=9872321;
RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA Goddard A.D., Botstein D., Ashkenazi A.;
RT "Genomic amplification of a decoy receptor for Fas ligand in lung and
RT colon cancer.";
RL Nature 396:699-703(1998).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-35.
RC TISSUE=Prostate;
RX MEDLINE=99253915; PubMed=10318773;
RA Yu K.-Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT "A newly identified member of tumor necrosis factor receptor
RT superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL J. Biol. Chem. 274:13733-13736(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=20122600; PubMed=10655513;
RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT "Overexpression of M68/Dcr3 in human gastrointestinal tract tumors
RT independent of gene amplification and its location in a four-gene
RT cluster.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
RN [4]
RP SEQUENCE FROM N.A.


```
FT REPEAT      83      125      TNFR-CYS 2.
FT REPEAT      126      166
FT REPEAT      167      195      TNFR-CYS 3.
FT DOMAIN      340      350      TNFR-CYS 4.
FT DOMAIN      362      447      DEATH.
FT DISULFID     44      58      BY SIMILARITY.
FT DISULFID     59      72      BY SIMILARITY.
FT DISULFID     62      81      BY SIMILARITY.
FT DISULFID     84      99      BY SIMILARITY.
FT DISULFID    102     117      BY SIMILARITY.
FT DISULFID    105     125      BY SIMILARITY.
FT DISULFID    127     143      BY SIMILARITY.
FT DISULFID    146     158      BY SIMILARITY.
FT DISULFID    149     166      BY SIMILARITY.
FT DISULFID    168     179      BY SIMILARITY.
FT DISULFID    182     194      BY SIMILARITY.
FT DISULFID    185     190      BY SIMILARITY.
FT CARBOHYD     54      54      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD     86      86      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD    145     145      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD    151     151      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE    461 AA; 50696 MW; CD72361EC60C9D43 CRC64;

Query Match      8.5%; Score 176.5; DB 1; Length 461;
Best Local Similarity 22.3%; Pred. NO. 0.00075;
Matches 99; Conservative 41; Mismatches 131; Indels 173; Gaps 26;

QY 2 PODR-PFEDTC-HGNPSHYDKAVRCCYRCPMGLFPTQOC--PQRPDCRKQCEPDYVL 57
Db 34 PDREKRESLCPQGYHPQNSK---CCTKCHKGYLHNDCLGPGLDTCR-ECONGTFT 90
QY 58 DEADRCTACVTC--RDDL--VEKTPCAWNSSRVCECRPMFCFS--TSAVNSCARCFFHS 111
Db 91 ASENHLTQCLSCSKCRSEMSQVEISPCIVDRTVCGCRKNQYKRWSEFLQCLNC---S 147
QY 112 VCPAGMIWKFPGTAQNTVCEPASGV--SPACASPEMCKEPPSGTIPQAKTPVSPATS 169
Db 148 LCPNG-VQLPCLERKODTICNCHSGFFELRDKCVSCVNCK-----NADCKNLCPTS 198
QY 170 SAST-----MPVRGCTRLA-----QBAASKLR---APDPSPSVGRP 203
Db 199 ETRNDFQDTGTVLLPFIFFGLCLAFFLVGLACRYQRWKPKLYSIICGKSTPVKEGP 258
QY 204 -----SSDPGLSPT---OPCPGSGDKCKQCEPDYVLDAGRCTACVSCS 245
Db 259 EPLATAPSGPTITTSPIPSFSPTTFSPVPSFS-----PISSPTF----- 299
QY 246 RDDLVEKTPCAWNSRSTCECRPMGICATSATNSCARCVYPICAGETVTKPDMAEKDTT 305
Db 300 -----TPCDWSNIK-----VTSPPKAIA----- 317
QY 306 FEAPPLGTQPCNTPENGAPASTSPQSLVDSQASKTLPIPT-----SA 352
Db 318 ---PPQAGP-ILPMP-----PAST-----VPVTPPKWGSAAHSA 353
QY 353 PVALS-----STGKPVLDAGPVLFW 372
Db 354 PAQLADADPATLYAVDGVPPTRW 377

RESULT 12
MUA3_CAEEL
AC P34576; Q90A13; STANDARD; PRT; 3767 AA.
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transmembrane cell adhesion receptor mua-3 precursor.
GN MUA-3 OR K085.3/T20G5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lu Z., Vogel B., Hedgecock E.;
RT "mua-3 mRNA splicing pattern revealed.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Berks M., Smith A., Kershaw J.K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Involved in cell adhesion.
CC -!- SIMILARITY: CONTAINS 52 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 SEA DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF139060; AAD29428.1; -.
DR EMBL; Z30974; CAA83226.2; -.
DR EMBL; Z30423; CAA83226.2; JOINED.
DR EMBL; Z30423; CAC42345.1; -.
DR EMBL; Z30974; CAC42345.1; JOINED.
DR PIR; S42373; S42373.
DR HSSP; P17301; IA0X.
DR WormPep; K08E5.3; CE28049.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR000082; SEA_domain.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00008; EGF; 33.
DR Pfam; PF00057; ldl_recept_a; 3.
DR Pfam; PF01390; SEA; 2.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00001; EGF_like; 45.
DR SMART; SM00192; LDLA; 4.
DR SMART; SM00200; SEA; 2.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 32.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 6.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50068; LDLRA_2; 4.
DR PROSITE; PS50024; SEA; 2.
DR PROSITE; PS50234; VWFA; 2.
KW Receptor; Cell adhesion; EGF-like domain; Repeat; Transmembrane;
KW Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 3767
FT POTENTIAL.
FT TRANSMEMBRANE CELL ADHESION RECEPTOR
FT MUA-3.
FT DOMAIN 25 3417
FT DOMAIN 3418 3438
FT DOMAIN 3439 3767
FT CYTOPLASMIC (POTENTIAL).
FT LDL-RECEPTOR CLASS A 1.
FT LDL-RECEPTOR CLASS A 1.
FT LDL-RECEPTOR CLASS A 2.
FT LDL-RECEPTOR CLASS A 3.
FT LDL-RECEPTOR CLASS A 4.
FT EGF-LIKE 1.
FT EGF-LIKE 2.
FT DOMAIN 375 416
```

```
FT DOMAIN 418 466 EGF-LIKE 3. BY SIMILARITY. 712
FT DOMAIN 468 517 EGF-LIKE 4. BY SIMILARITY. 729
FT DOMAIN 519 566 EGF-LIKE 5. BY SIMILARITY. 738
FT DOMAIN 614 663 EGF-LIKE 6. BY SIMILARITY. 759
FT DOMAIN 665 713 EGF-LIKE 7. BY SIMILARITY. 779
FT DOMAIN 714 760 EGF-LIKE 8. BY SIMILARITY. 788
FT DOMAIN 762 810 EGF-LIKE 9. BY SIMILARITY. 809
FT DOMAIN 816 860 EGF-LIKE 10. BY SIMILARITY. 836
FT DOMAIN 861 908 EGF-LIKE 11. BY SIMILARITY. 845
FT DOMAIN 910 961 EGF-LIKE 12. BY SIMILARITY. 859
FT DOMAIN 963 1012 EGF-LIKE 13. BY SIMILARITY. 879
FT DOMAIN 1029 1070 EGF-LIKE 14. BY SIMILARITY. 888
FT DOMAIN 1071 1118 EGF-LIKE 15. BY SIMILARITY. 907
FT DOMAIN 1120 1168 EGF-LIKE 16. BY SIMILARITY. 930
FT DOMAIN 1170 1219 EGF-LIKE 17. BY SIMILARITY. 941
FT DOMAIN 1230 1406 WFA. BY SIMILARITY. 967
FT DOMAIN 1421 1466 EGF-LIKE 18. BY SIMILARITY. 981
FT DOMAIN 1466 1510 EGF-LIKE 19. BY SIMILARITY. 990
FT DOMAIN 1521 1562 EGF-LIKE 20. BY SIMILARITY. 1011
FT DOMAIN 1563 1608 EGF-LIKE 21. BY SIMILARITY. 1033
FT DOMAIN 1608 1656 EGF-LIKE 22. BY SIMILARITY. 1046
FT DOMAIN 1658 1706 EGF-LIKE 23. BY SIMILARITY. 1055
FT DOMAIN 1708 1755 EGF-LIKE 24. BY SIMILARITY. 1069
FT DOMAIN 1759 1807 EGF-LIKE 25. BY SIMILARITY. 1087
FT DOMAIN 1809 1860 EGF-LIKE 26. BY SIMILARITY. 1096
FT DOMAIN 1862 1911 EGF-LIKE 27. BY SIMILARITY. 1117
FT DOMAIN 1913 1961 EGF-LIKE 28. BY SIMILARITY. 1137
FT DOMAIN 1963 2011 EGF-LIKE 29. BY SIMILARITY. 1146
FT DOMAIN 2014 2062 EGF-LIKE 30. BY SIMILARITY. 1167
FT DOMAIN 2068 2112 EGF-LIKE 31. BY SIMILARITY. 1188
FT DOMAIN 2113 2160 EGF-LIKE 32. BY SIMILARITY. 1197
FT DOMAIN 2162 2208 EGF-LIKE 33. BY SIMILARITY. 1218
FT DOMAIN 2210 2258 EGF-LIKE 34. BY SIMILARITY. 1241
FT DOMAIN 2260 2308 EGF-LIKE 35. BY SIMILARITY. 1433
FT DOMAIN 2310 2358 EGF-LIKE 36. BY SIMILARITY. 1450
FT DOMAIN 2360 2408 EGF-LIKE 37. BY SIMILARITY. 1465
FT DOMAIN 2409 2455 EGF-LIKE 38. BY SIMILARITY. 1484
FT DOMAIN 2456 2504 EGF-LIKE 39. BY SIMILARITY. 1494
FT DOMAIN 2513 2563 EGF-LIKE 40. BY SIMILARITY. 1478
FT DOMAIN 2565 2616 EGF-LIKE 41. BY SIMILARITY. 1496
FT DOMAIN 2618 2666 EGF-LIKE 42. BY SIMILARITY. 1509
FT DOMAIN 2668 2714 EGF-LIKE 43. BY SIMILARITY. 1525
FT DOMAIN 2716 2763 EGF-LIKE 44. BY SIMILARITY. 1532
FT DOMAIN 2763 2811 EGF-LIKE 45. BY SIMILARITY. 1547
FT DOMAIN 2811 2872 EGF-LIKE 46. BY SIMILARITY. 1561
FT DOMAIN 2873 2977 SEA 1. BY SIMILARITY. 1583
FT DOMAIN 2977 3048 EGF-LIKE 47. BY SIMILARITY. 1592
FT DOMAIN 3049 3172 SEA 2. BY SIMILARITY. 1594
FT DOMAIN 3172 3220 EGF-LIKE 48. BY SIMILARITY.
FT DOMAIN 3224 3272 EGF-LIKE 49. BY SIMILARITY.
FT DOMAIN 3272 3324 EGF-LIKE 50. BY SIMILARITY.
FT DOMAIN 3328 3373 EGF-LIKE 51. BY SIMILARITY.
FT DOMAIN 3373 3409 EGF-LIKE 52. BY SIMILARITY.
FT DISULFID 229 243 BY SIMILARITY.
FT DISULFID 235 252 BY SIMILARITY.
FT DISULFID 254 267 BY SIMILARITY.
FT DISULFID 381 392 BY SIMILARITY.
FT DISULFID 386 402 BY SIMILARITY.
FT DISULFID 404 415 BY SIMILARITY.
FT DISULFID 422 435 BY SIMILARITY.
FT DISULFID 429 444 BY SIMILARITY.
FT DISULFID 446 465 BY SIMILARITY.
FT DISULFID 472 486 BY SIMILARITY.
FT DISULFID 480 495 BY SIMILARITY.
FT DISULFID 497 516 BY SIMILARITY.
FT DISULFID 523 536 BY SIMILARITY.
FT DISULFID 530 545 BY SIMILARITY.
FT DISULFID 547 565 BY SIMILARITY.
FT DISULFID 618 632 BY SIMILARITY.
FT DISULFID 626 642 BY SIMILARITY.
FT DISULFID 644 662 BY SIMILARITY.
FT DISULFID 669 682 BY SIMILARITY.
FT DISULFID 676 691 BY SIMILARITY.

Query Match 8.5%; Score 176; DB 1; Length 3767;
Best Local Similarity 22.2%; Pred. No. 0.0049;
Matches 91; Conservative 44; Mismatches 162; Indels 112; Gaps 21;

QY 6 PFEDTCHGNPSHYDKAVRRCCYRCMPG---LFTPTQO---CPQRPDCKRQCEPDYLD 58
Db 868 PLRNTC-GKNAECIDLAEGYTC-QCRSGYADISPVQPGRICRARNVCSN--KEKYNVD 923
QY 59 EADRCTACTVCSRDDLVKPTCAWNSRVCECRPGMFCSTSAVNSCARCFHFHVCVPAGMI 118
Db 924 -----CSENAICADTEHSYS---CRCRPGFADVSAAFN----- 953
QY 119 VKFPFGTAQNTVCEPASPVGSPACASPCNKPSGSGTIPQAKPTVPSPATSSASTMPVRG 178
Db 954 -KLPGRRCTEAVNECASPSLND-CSKNAFCEDAKEGVICTCRPGYVDNSPNA-----RH 1006
QY 179 GTRLAQEAASKL-TRAPDSPSSVGRPSDFGLSPTOP-----CPEGSGDCRQCEPDY 230
Db 1007 PGRICTKPVEKIKTDLKDTSF-----STDGCDPKPKGCAGNEACVQRHGHQHCVEYETA 1061
QY 231 YLDEAGRCTACVSCSRDDLVEKTPCAWN--SSRTCECRPGMI-----CATSAT 276
Db 1062 FRYTDGSCRVSACSRRNTCDKNAICLNRFDSYTCQCRPGYIDLSDALT/NAPGRICKELI 1121
QY 277 NSCA-----RCVPYPICA---GETVTKPQDMAEKDTFEAPP-----LGT 313
Db 1122 NECAASDNECSPYARCIDATNGYACQCLDGFIDVSSRYNKKPGRQCTNSNNECEKSLNT 1181
QY 314 ---QPCDNPTPENGAPASTSPQTQSLLDVDSQASKTLP-----IPTSAP 353
```

Db 1182 CDENABCDVTPD-----GYTCQCYGGEVDVSSNANLPGRVCTVQTCP 1225

RESULT 13

```

TR21_MOUSE
ID TR21_MOUSE STANDARD; PRT; 655 AA.
AC Q9EP05; Q91XH9; Q91W77;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 21 precursor (TNFR-
DE related death receptor-6) (Death receptor 6).
OS TNFRSF21 OR DR6.
OS Mus musculus (Mouse).
OC STRAIN=C57BL/6; TISSUE=Kidney;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Kidney;
RA Isogai D., Ichino M., Yoshinari M., Yamaura A., Kurokawa F.,
RA Minami M.;
RT "Mouse DR6: mouse homolog of human TNFR-related death receptor-6
RT (DR6)".
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Kidney;
RA Kim V., Machleidt T., Shi W.-X., Wang X., Cai Z.;
RA "Murine DR6: murine TNFR-related death receptor-6.";
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RN FUNCTION
RX MEDLINE=21571606; PubMed=11714751;
RA Zhao H., Yan M., Wang H., Erickson S., Grewal I.S., Dixit V.M.;
RT "Impaired c-Jun amino terminal kinase activity and T cell
RT differentiation in death receptor 6-deficient mice.";
RL J. Exp. Med. 194:1441-1448(2001).
CC -!- FUNCTION: May activate NF-kappa-B and promote apoptosis (By
CC similarity). May activate JNK and be involved in T-cell
CC differentiation.
CC -!- FUNCTION: May activate NF-kappa-B and JNK and promote apoptosis.
CC May be involved in T-cell differentiation.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in adult spleen,
CC thymus, testis, prostate, ovary, small intestine, colon, brain,
CC lung and kidney, and in fetal brain, liver and lung. Detected at
CC lower levels in adult peripheral blood leukocytes, lung, and in
CC fetal muscle, heart, kidney, small intestine and skin. Detected in
CC T-cells, B-cells and monocytes. In T-cells expression is highest
CC in Th0 cells, intermediate in Th2 cells and lower in Th1 cells.
CC -!- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-25 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF322069; AAG38115.1; -
CC EMBL; AY043489; AAK74193.1; -
CC EMBL; BC016420; AAH16420.1; -
CC HSP; O14763; LD06.

```

```

DR MGD; MGI:2151075; Tnfrsf21.
DR InterPro; IPR000488; Death.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 4.
DR ProDom; PD000771; TNFR_c6; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00208; TNFR; 4.
DR PROSITE; PS00017; DEATH_DOMAIN; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
DR PROSITE; PS00050; TNFR_NGFR_2; 1.
KW Receptor; Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 41 POTENTIAL.
FT CHAIN 42 655 TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER 21.
FT DOMAIN 42 349 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 350 370 POTENTIAL.
FT DOMAIN 371 655 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 415 498 DEATH.
FT REPEAT 50 88 TNFR-CYS 1.
FT REPEAT 90 131 TNFR-CYS 2.
FT REPEAT 133 167 TNFR-CYS 3.
FT REPEAT 170 211 TNFR-CYS 4.
FT DISULFID 67 80 BY SIMILARITY.
FT DISULFID 70 88 BY SIMILARITY.
FT DISULFID 91 106 BY SIMILARITY.
FT DISULFID 109 123 BY SIMILARITY.
FT DISULFID 113 131 BY SIMILARITY.
FT DISULFID 133 144 BY SIMILARITY.
FT DISULFID 150 168 BY SIMILARITY.
FT DISULFID 171 186 BY SIMILARITY.
FT DISULFID 192 211 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 352 352 W -> L (IN REF. 1).
FT CONFLICT 523 523 M -> I (IN REF. 3).
SQ SEQUENCE 655 AA; 71982 MW; 5EC7C51C7C99EFF7 CRC64;

Query Match 8.4%; Score 174.5; DB 1; Length 655;
Best Local Similarity 22.6%; Pred. No. 0.0013;
Matches 53; Conservative 29; Mismatches 102; Indels 51; Gaps 6;

QY 27 CYRCPMGLFPTQCPQRPDCKRQCEPDYILDEADRCTACVTCSDR---DLVEKTPCAWN 83
DB 67 CDKCPAGTYVSEHCTNMSLRVCSQCPACTFTRHENGIERCHDCSQCPWPMIERLPCAAL 126
QY 84 SSRVCECPGMCFTSAVNSCARCFHFSVCPAGMIVKFPGTAQKNTVCEPASPG----- 137
DB 127 TDREICPPGMYSQNGT-----CAPHTVCPVGWGVRRKGTENEDVRCKQCARGFSDVP 180
QY 138 -----VSPACASPEN-----CKPESSGTTPQAQKPTVSPAT 168
DB 181 SSVMKCAHDTCLQNLNVKPGKEDTVNCGMLFFSTNPPSSGTFTFSPEHMSHD 240
QY 169 SSATMEVRVG-GLRLAQEAASAKLTPADS-----PSSVGRFPSSDPGLSPTQCP 216
DB 241 VPSSTYEPQGMNSTDNSASTASVRKVPSSGIEGTVPDNTSSTSGKEGTNRTLPNP 295

RESULT 14
PK5_BRACL
ID PK5_BRACL STANDARD; PRT; 1696 AA.
AC Q9NJ15; Q9NJ16; Q9NJ14;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
DE (Protein convertase PC6-like) (apc6).
GN PC6.
OS Branchiostoma californiensis (California lancelet) (Amphioxus).

```


CC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OX Branchiostoma.
 CC NCBI_TaxID=7738;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
 RX MEDLINE=20175281; PubMed=10708868;
 RA Oliva A.A. Jr., Chan S.J., Steiner D.F.;
 RT "Evolution of the prohormone convertases: identification of a
 RL Biochim. Biophys. Acta 1477:338-348(2000).
 CC -!- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY
 CC OF CLEAVAGE AT THE RX(R/R)R CONSENSUS MOTIF (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
 CC PROTEINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-|-ZAA BONDS, WHERE XAA
 CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
 CC -!- SUBCELLULAR LOCATION: ISOFORM A AND ISOFORM C ARE SECRETED.
 CC ISOFORM B IS A TYPE I MEMBRANE PROTEIN.
 CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A, B (SHOWN HERE) AND C; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
 CC ASSISTING THE FOLDING OF THE ZYMOGEN WITHIN THE ENDOPLASMIC
 CC RETICULUM.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 CC -!- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF184615; AAF26300.1; -;
 DR EMBL; AF184616; AAF26301.1; -;
 DR EMBL; AF184617; AAF26302.1; -;
 DR HSSP; Q99405; IMPT.
 DR MEROPS; S08.UPB; -;
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR002174; Furin-like.
 DR InterPro; IPR002884; P_domain.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF01483; P; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR ProDom; PD000717; P_domain; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00261; FU; 17.
 DR PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
 KW Cleavage on pair of basic residues; Repeat; Alternative splicing;
 KW Transmembrane.
 FT SIGNAL 1 25 POTENTIAL.
 FT PROPEP 26 110 POTENTIAL.
 FT CHAIN 111 1696 PROTEIN CONVERTASE SUBTILISIN/KEXIN
 FT TYPE 5.
 FT DOMAIN 111 1618 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1619 1639 POTENTIAL.
 FT DOMAIN 1640 1696 POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT CATALYTIC.
 FT HOMO B.
 FT DOMAIN 496 637 HOMO B.
 FT DOMAIN 664 1649 CYS-RICH MOTIF (CRM) REGION.
 FT SITE 110 111 CLEAVAGE (AUTO-) (BY SIMILARITY).
 FT ACT_SITE 192 192 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 233 233 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 407 407 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 529 529 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 885 885 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 1259 1323 DDTILDRGECITSGPGGYMDRREKKCKKACHTCKECSDEY

FT DTCTACNDGFLLTASSCEAGCP -> AENQNASCPFFA
 FT PREVSVAELALGHLRYSLTDVPPQSNPPDTVLGADRRL
 FT TTATSAARGCA (IN ISOFORM C).
 FT MISSING (IN ISOFORM C).
 FT VARSPLIC 1324 1696
 FT VARSPLIC 1288 1343
 FT CHTPKCSEYDDDTCTACNDGFLLTASSCEAGCPGQFL
 FT HGGDCDSCHRECKTC -> IARCVDRLDRSCDLVLRNFC
 FT VRRIFVRCCTGCKLYWEDRPMRRGSSOPTQGRN (IN
 FT ISOFORM A).
 FT VARSPLIC 1344 1696
 FT MISSING (IN ISOFORM A).
 SQ SEQUENCE 1696 AA; 188410 MW; 281CBEL784257CBD CRC64;
 Query Match 8.3%; Score 172.5; DB 1; Length 1696;
 Best Local Similarity 17.2%; Pred. No. 0.0039;
 Matches 99; Conservative 51; Mismatches 161; Indels 265; Gaps 26;
 QY 10 TCHGNPS-----HYDKAVRR-----CYRCPMGLEPT 37
 DB 686 TCHGRISGQCLSCHEGNYFVEDEGTCEECGQYKDEERKCLDCSADCLTCQVSADHC 745
 QY 38 QQCPQR-----PTDCRKQCEPDYLLDEADRC-----TACVTCSDLLVE 76
 DB 746 TSCDDEGLKLFENTCVAQCSEGRYMDNDVCQDDSDCTCTGPDATDCVTCADLILQ 805
 QY 77 KTPCAWNSR-----
 DB 806 ESQCVCESGSGYFQOEYELCKHATCASCSGSRDDQCLTCSGHLELDETHRCITSCBDG 865
 QY 87 -----VC-----ECRP-- 92
 DB 866 EYTERGKCEDCNICKKNGSQADQCLEHHDNLNLYDTTCVQCGNRRYPENGECHPCH 925
 QY 93 --GMFCSTSAVNSCARF-----FHSVCPAGMIVKFPGTAKNTVCEPASP 136
 DB 926 PSLGICIGGIEINOCNICTDYEGEDHFLYQGTCHVTCTPPGL---YGDITDQ--VCKACAP 980
 QY 137 GVSPACASPEN-----CKE---PSSGTIPQAKPTVPSPATSSASTMPVRGT----- 180
 DB 981 GCI-ACDGPADNQTLCEERAPTDRGRC-QSEGSOTDEAECAEGCHSCEGPDICDSCE 1038
 QY 181 --RLAEEAASKLTRA-----PDSFSSVGRPSD-----PGLSPTQPCPG-----SGDC 222
 DB 1039 DYLTEDTCVRRNCPSTFTYDDODRECRCHDNCACDGNPNQNCCKEYKFTPDGC 1098
 QY 223 RKQCEPDYLLDEAGR-----CTA-----CVSCSRDDLVEKTPC----- 255
 DB 1099 STCCPNRYVKDDFNKECKPCDSCFTCSGSPASPHCLSCADGDFLHSSCRCTCPAGTGN 1158
 QY 256 -----AWNRSRT--CECRPGMICATSNASCARCP--YPICAGETVTKPDQ 298
 DB 1159 AESHECVSSCEQDQYYSSETGRCEDCPYNCACDNDGDCAECACTVIVDGR--CRPEE 1216
 QY 299 MAE---KDTTFEAPLGTQPCNPTPENGEPAST 330
 DB 1217 TCEDGEYQDRDRDTAELSCR-CHQSCKTCGSPSDT 1251
 RESULT 15
 ZAN_MOUSE
 ID ZAN_MOUSE STANDARD; PRT; 5376 AA.
 AC O88799; O08647;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Zonadhesin precursor.
 GN ZAN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=98123114; PubMed=9452463;

RA Gao Z., Garbers D.L.;
 RT "Species diversity in the structure of zonadhesin, a sperm-specific
 RT membrane protein containing multiple cell adhesion molecule-like
 RT domains.";
 RT J. Biol. Chem. 273:3415-3421(1998).
 RN [2]
 RP SEQUENCE OF 4864-5376 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=97271566; PubMed=9126492;
 RA Gao Z., Harumi T., Garbers D.L.;
 RT "Chromosome localization of the mouse zonadhesin gene and the human
 RT zonadhesin gene (ZAN).";
 RL Genomics 41:119-122(1997).
 CC -!- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
 CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
 CC SIGNALING.
 CC -!- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
 CC APICAL REGION OF THE SPERM HEAD.
 CC -!- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
 CC -!- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATES SPERM ADHESION TO THE
 CC ZONA PELLUCIDA.
 CC -!- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
 CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
 CC SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
 CC -!- DOMAIN: THE WVF DOMAIN 2 MAY MEDIATE COVALENT
 CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
 CC -!- SIMILARITY: CONTAINS 3 MAM DOMAINS.
 CC -!- SIMILARITY: CONTAINS 25 WVF DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U97068; AAC26680.1; -;
 DR EMBL; U83190; AAC53125.1; -;
 DR MGD; MGI:106656; Zan.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR003645; FoIn.
 DR InterPro; IPR000998; MAM_domain.
 DR InterPro; IPR002919; TIL_Cysrich.
 DR InterPro; IPR003328; Tila_Cysrich.
 DR InterPro; IPR001007; VWF_C.
 DR InterPro; IPR001846; VWF_D.
 DR Pfam; PF00094; vwd; 4.
 DR Pfam; PF00629; MAM; 3.
 DR Pfam; PF01822; TIL; 25.
 DR Pfam; PF02345; Tila; 25.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00274; FOLN; 11.
 DR SMART; SM00137; MAM; 2.
 DR SMART; SM00214; VWC; 17.
 DR SMART; SM00216; VWD; 4.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 18.
 DR PROSITE; PS00740; MAM_1; FALSE_NEG.
 DR PROSITE; PS50060; MAM_2; 3.
 KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
 KW Repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 5376
 FT DOMAIN 18 5310 ZONADHESIN.
 FT TRANSMEM 5311 5337 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 5338 5376 POTENTIAL.
 FT DOMAIN 45 210 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 215 374 MAM 1.
 FT DOMAIN 377 542 MAM 2.
 FT DOMAIN 547 1170 MAM 3.
 FT 80 X HEPTAPEPTIDE REPEATS (APPROXIMATE)

FT DOMAIN 1171 1280 (MUCIN-LIKE DOMAIN).
 FT WVF1 1 (PARTIAL).
 FT WVF2 2 (PARTIAL).
 FT WVF3 3.
 FT WVF4 4.
 FT WVF5 5 (PARTIAL).
 FT WVF6 6 (PARTIAL).
 FT WVF7 7 (PARTIAL).
 FT WVF8 8 (PARTIAL).
 FT WVF9 9 (PARTIAL).
 FT WVF10 10 (PARTIAL).
 FT WVF11 11 (PARTIAL).
 FT WVF12 12 (PARTIAL).
 FT WVF13 13 (PARTIAL).
 FT WVF14 14 (PARTIAL).
 FT WVF15 15 (PARTIAL).
 FT WVF16 16 (PARTIAL).
 FT WVF17 17 (PARTIAL).
 FT WVF18 18 (PARTIAL).
 FT WVF19 19 (PARTIAL).
 FT WVF20 20 (PARTIAL).
 FT WVF21 21 (PARTIAL).
 FT WVF22 22 (PARTIAL).
 FT WVF23 23 (PARTIAL).
 FT WVF24 24 (PARTIAL).
 FT WVF25.
 FT EGF-LIKE.
 FT BY SIMILARITY.
 FT DISULFID 5263 5274 BY SIMILARITY.
 FT DISULFID 5268 5283 BY SIMILARITY.
 FT DISULFID 5285 5294 BY SIMILARITY.
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1216 1216 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1239 1239 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1314 1314 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1814 1814 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1908 1908 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1933 1933 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2028 2028 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2111 2111 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2142 2142 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2332 2332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2533 2533 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2575 2575 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2692 2692 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2812 2812 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 3052 3052 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 3065 3065 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 3144 3144 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 3172 3172 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 3288 3288 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 3292 3292 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 3782 3782 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 4005 4005 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 4136 4136 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 4243 4243 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 4254 4254 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 4335 4335 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 4376 4376 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 4586 4586 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 5136 5136 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 5252 5252 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 5376 AA; 579908 MW; 0E44DB77DF2A2620 CRC64;

Query Match 8.2%; Score 171.5; DB 1; Length 5376;
 Best Local Similarity 22.9%; Pred. No. 0.012;
 Matches 89; Conservative 34; Mismatches 127; Indels 139; Gaps 25;
 QY 16 SHYDKAVRCCYRC-PMGLFPTQCC-----PQRTDGRK--QCEPDYVLEADRCTACTV 68
 Db 3781 SNYTD-----CLPFCPLPCLDPSALCGTSPKGPSTCKEGVCQPGVLDK-DKCIKIE 3834
 QY 69 CS-RDDLVEKTPC--AWNS-----SRVC-----ECRPGMFCSTSAVNSCAR----- 106

```
Db 3835 CGCRDTQGAVIPAGKTWLTGCIQSCACVEGTIQCNFQCPPEGTYNHN--NNCAKIPLQ 3892
Qy 107 CFFHS-----VCPAGMIVKFPGTAKNTVCEPASPQVSPAC-----ASPE----- 146
Db 3893 CPAHSHFTSCLPSCF-----PSCANLDGSCQETSPKVPSTCKEGCLCOPGYFLNNGKC 3945
Qy 147 -----NCKEFPSSGTIPQAKPTVPSPATSSASTMPVRGGTRLAQ-----EA 186
Db 3946 VLOTHCDCKDAEGGLVPAGKTWTSKDTQSCA---CTGGAVOCQNFQCPPLGTICKDSGDG 4002
Qy 187 ASKLT-----RAPDSFSSVGRPSDDPCLSPQPCPEGGDCRQQ---- 225
Db 4003 SSNCTKIHKGAMGCVLMAGGIRALQCPAHSHFTSCLPSCPPS--CSNLDGSCVESNFKA 4060
Qy 226 -----CEPDYVLDAGRCTACVSCS-RDDLVEKTPC--AWNSS----- 260
Db 4061 PSVCKKGCICOPGYLLNN-DKCVLRIQCCKDTQGLIPAGRTWISSDCTKSCSCMGGII 4119
Qy 261 --RTCECRPGMICATNSCARCVPI 287
Db 4120 QCRDFQCPPTGYCKES-NDSSRTCAKIPL 4147
```

Search completed: December 11, 2002, 15:24:22
Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:22:16 ; Search time 22 Seconds
(without alignments)

1625.546 Million cell updates/sec

Title: US-09-628-126-2_COPY_19_390

Perfect score: 2080

Sequence: 1 FQDRPFDTCHGNPSHYD.....PVALSSTGKPVLDAGPVLFW 372

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2074	99.7	595	2	A42086	CD30 antigen precursor
2	618	29.7	493	2	J05486	membrane glycoprotein
3	225.5	10.8	461	1	A35356	tumor necrosis fac
4	216	10.4	435	2	I54182	tumor necrosis fac
5	200.5	9.6	474	2	B38634	tumor necrosis fac
6	198.5	9.5	459	2	I48854	gene murine tumour
7	198.5	9.5	1299	2	T43251	furin (EC 3.4.21.7
8	176.5	8.5	461	2	J04302	tumor necrosis fac
9	176	8.5	3051	2	S42373	hypothetical prote
10	171.5	8.2	5376	2	T42215	zonadhesin - mouse
11	171	8.2	1797	2	A55677	laminin beta-2 cha
12	166	8.0	1820	2	A55494	latent transformin
13	165.5	8.0	1274	2	T42017	cysteine rich prot
14	165.5	8.0	1428	2	T08852	lustrin A - Califo
15	165.5	8.0	1798	2	S53869	laminin beta-2 cha
16	165	7.9	2321	2	S78549	notch3 protein - h
17	165	7.9	2718	2	A23475	G surface protein
18	164	7.9	325	2	B43692	T2 protein - rabbi
19	161.5	7.8	572	2	T29880	hypothetical prote
20	161.5	7.8	802	2	T24293	hypothetical prote
21	161.5	7.8	949	2	T24294	hypothetical prote
22	161.5	7.8	1766	2	A42125	trophozoite cystei
23	159	7.6	2871	2	A55567	fibrillin I - bovi
24	159	7.6	3570	2	T45025	mucin MUC5B, trach
25	158	7.6	1574	2	T13994	MEGF6 protein - ra
26	158	7.6	2704	2	S09118	G surface protein
27	157.5	7.6	1680	2	A43434	furin (EC 3.4.21.7
28	156.5	7.5	1955	1	AGCH	agrin precursor -
29	154.5	7.4	327	2	A46484	apoptosis-mediatin

ALIGNMENTS

RESULT 1

A42086

CD30 antigen precursor - human

N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000

C:Accession: A42086

R:Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.

Cell 68, 421-427, 1992

A:Title: Molecular cloning and expression of a new member of the nerve growth factor

A:Reference number: A42086; MUID:92154659; PMID:1310894

A:Accession: A42086

A:Molecule type: mRNA

A:Residues: 1-595 <DUR>

A:Cross-references: GB:M83554; NID:g180095; PIDN:AAA51947.1; PID:g180096

A:Experimental source: HUM-102 cell line

A:Note: sequence extracted from NCBI backbone (NCBIN:82088, NCBI:P:82090)

C:Genetics:

A:Gene: GDB:CD30; DLS166E

A:Cross-references: GDB:131547; OMIM:153243

A:Map position: lp36-lp36

C:Superfamily: NGF receptor repeat homology

C:Keywords: glycoprotein; growth factor receptor; transmembrane protein

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-383/Domain: extracellular #status predicted <EXT>

F:384-407/Domain: transmembrane #status predicted <TM>

F:408-595/Domain: intracellular #status predicted <CYT>

F:101.276/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.7%; Score 2074; DB 2; Length 595;

Best Local Similarity 99.7%; Pred. No. 5e-109;

Matches 371; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQDRPFDTCHGNPSHYDVKAVRCCYRCPCMGFLPTQCPQRPDTCRCKQCEPDYILDEA 60

Db 19 FQDRPFDTCHGNPSHYDVKAVRCCYRCPCMGFLPTQCPQRPDTCRCKQCEPDYILDEA 78

Qy 61 DRCTACVTCSDLDLVEKTPCAWNSRVCRCRPGMFCSTAVNSCARCFFHSVCAPGAMIVK 120

Db 79 DRCTACVTCSDLDLVEKTPCAWNSRVCRCRPGMFCSTAVNSCARCFFHSVCAPGAMIVK 138

Qy 121 FPGTAQNTVCEPASPGVSPACAPENCKEPSSGTIPQAKPTVPVSPATSSASTMPVRGTT 180

Db 139 FPGTAQNTVCEPASPGVSPACAPENCKEPSSGTIPQAKPTVPVSPATSSASTMPVRGTT 198

Qy 181 RLAAQEAASLKLTRAPDSPSSVGRSPDGLSPTQPCPGSGDCRCKQCEPDYILDEAGRCTA 240

Db 199 RLAAQEAASLKLTRAPDSPSSVGRSPDGLSPTQPCPGSGDCRCKQCEPDYILDEAGRCTA 258

Qy 241 CVCSCSDLDLVEKTPCAWNSRVCRCRPGMFCSTAVNSCARCVPYPICAGETVTKPDMDA 300

Db 259 CVCSCSDLDLVEKTPCAWNSRVCRCRPGMFCSTAVNSCARCVPYPICAAETVTKPDMDA 318

notch 3 protein -
hypothetical prote
gene G4R protein -
hypothetical prote
mucin 2 precursor,
G2R protein - vari
fibrillin 1 precur
hypothetical prote
laminin alpha 5 ch
laminin beta-2 cha
death receptor-6 -
T2 protein - myxom
tumor necrosis fac
variant-specific s
gastric mucin MUC5
glucan 1,4-alpha-9

30 154.5 7.4 2318 2 S45306
31 153.5 7.4 348 2 T28623
32 153.5 7.4 349 2 D36858
33 153.5 7.4 839 2 F75518
34 153 7.4 3020 2 A43932
35 152.5 7.3 349 2 D72175
36 152 7.3 3002 2 A47221
37 151.5 7.3 728 2 T20561
38 151.5 7.3 3635 2 T10053
39 151 7.3 1801 1 MMRTS
40 150.5 7.2 651 2 JC7705
41 150 7.2 326 1 GQVZML
42 150 7.2 455 1 GQUT1
43 149.5 7.2 557 2 A48434
44 148 7.1 850 2 S56015
45 148 7.1 1367 1 S48478

F:262-279/Domain: transmembrane #status predicted <TMN>
F:280-461/Domain: intracellular #status predicted <INT>
F:171,193/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.8%; Score 225.5; DB 1; Length 461;
Best Local Similarity 28.0%; Pred. No. 6.5e-06;
Matches 61; Conservative 24; Mismatches 114; Indels 19; Gaps 6;

QY 18 YDKAVRRCYRCPCMGILFTQCCPQRPPTDCRKOCEPDYLL---DEADRCFACVTCRSDDL 74
Db 45 YYDQTAQMCCKSKCPQHAKVCTKTSDTVCDSEDSYTLQWVWPECLSCGRSSDQ 104
QY 75 VEKTPCAWNSRVCECPGMFCSTSAVNSCARCFHFSVCPAGMIVKPPGTAQKNTVCEPA 134
Db 105 VETQACTREQNRICTRCPGWYCALSKQEGCRLCAPLKRCPGFGVARPGTETSDVWCKPC 164
QY 135 SPQV-SPACASPNCNKEPSSGIIPQAKPTVPVSPATSSATMPVYR----GGTRLAQRAASK 189
Db 165 APGTFTNTSSTDICR-PHQICNVVAIPGNASMDAVCTSTSPTRMAPGAVHLPQPVSTR 223
QY 190 LTRAPDPSVSGVRPSDPLGLS-----PTQCPCEGSGD 221
Db 224 SQHTQPTP----EPSTAPSTFLLPMGPSPAPBAGSTGD 257

RESULT 4

I54182

tumor necrosis factor receptor 2-related protein - human

C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 17-Mar-2000
C:Accession: I54182
R:Baens, M.; Chaffanet, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.
Genomics 16, 214-218, 1993

A:Title: Construction and evaluation of a hncDNA library of human 12p transcribed sequences
A:Reference number: I54182; MUID:93252381; PMID:8486360
A:Accession: I54182

A:Status: preliminary;

A:Molecule type: mRNA

A:Residues: 1-435 <RES>

A:Cross-references: GB:I04270; NID:g339761; PIDN:AAA36757.1; PID:g339762

C:Genetics:

A:Gene: GDB:LTBR

A:Cross-references: GDB:I230195; OMIM:600979

A:Map position: 12p13.3-12p13.1

C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

Query Match 10.4%; Score 216; DB 2; Length 435;
Best Local Similarity 23.2%; Pred. No. 2.1e-05;
Matches 82; Conservative 30; Mismatches 112; Indels 130; Gaps 14;

QY 10 TCHGNPSHYDKAVRCCYRCPMGLFPTQCCPO-RPTDCRKQCEPDYLLDEADRCTACTVT 68
Db 42 TCRDQKEYEYEPQHRITCCSRCPGTYVSAKCSRIQRTVC-ATCAENSYNEHNYLTICQL 100
QY 69 CSRDD---LVKPTPCAWNSRVCECPGMFCSTSAV----- 101
Db 101 CRPCDPMVLEETAPCTSRKTCRCQCPGMFCMAWLECTHCELLSDCPGTEAELKDEV 160
QY 102 ----NSC-----ARCFHFSVCP-AGMIVKPPGTAQKNTVC-----EPASPG 137
Db 161 GKGNHCVPCKAGHFONTSSPASCOPHTRCENQGLVEAPAGTAQSDTTCKNPLELPPE 220
QY 138 VSP-----AC-----ASPENCKE-----PSSGTIPQ 158
Db 221 MSCTMLMLAVLLAPFLLLATVFCIMKSHPSLCRLKGLSLLKRRPQGEQPNPVAGSWEP 280
QY 159 AKPTP-----VSPATSSATMPVVRGTRLAQEAASKLTRAPD----SPSS 199
Db 281 PRAHPTFDLVPPLPIISGDVSPVSTGLPAAPVLEAGVQQQSPDLTLTREPQLEPEQSQ 340
QY 200 VGRPSSDPL-----SPTQPCPEGSGDCRKOCEPDYLLDEAG 236
Db 341 VAHGTN--GIHVTGSGMTITGNIYINGVPLGCPGPGDLPATPEPPYPIPEG 392

RESULT 5

B38634

tumor necrosis factor receptor type 2 precursor - mouse

C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C:Accession: B38634; A40254; S54816

R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen, Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991
A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis factor receptors
A:Reference number: A38634; MUID:91187885; PMID:1849278

A:Accession: B38634

A:Molecule type: mRNA

A:Residues: 1-474 <LEW>

A:Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828

R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; J Mol. Cell. Biol. 11, 3020-3026, 1991

A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors f

A:Reference number: A40254; MUID:91246168; PMID:1645445

A:Accession: A40254

A:Molecule type: mRNA

A:Residues: 1-474 <GOO>

A:Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828

R:Kissonerghis, M.; Fellowes, R.; Feldmann, M.; Chernaiovsky, Y. Submitted to the EMBL Data Library, May 1995

A:Description: Characterization of the promoter region of the murine p75-TNF receptor

A:Reference number: S54816

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-22 <KTS>

A:Cross-references: EMBL:X87128; NID:g809043; PIDN:CAA60618.1; PID:g809044

C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology

F:1-22/Domain: cytochrome receptor; transmembrane protein

F:23-474/Product: signal sequence #status predicted <SIG>

F:40-77/Domain: tumor necrosis factor receptor type 2 #status predicted <MAT>

F:79-120/Domain: NGF receptor repeat homology <NG1>

F:166-203/Domain: NGF receptor repeat homology <NG4>

Query Match 9.6%; Score 200.5; DB 2; Length 474;
Best Local Similarity 20.5%; Pred. No. 0.00017;
Matches 93; Conservative 39; Mismatches 160; Indels 161; Gaps 16;

QY 18 YDKAVRRCYRCPCMGILFPTQCCPQRPPTDCRKQCEPDYLLDEADRCTACTVTCSRDL 74
Db 46 YDKRAQMCCKAKCPGQYVVKHFNKTSDTVACDCEASMTQVWVNFRTCLSCSSCTTDQ 105
QY 75 VEKTPCAWNSRVCECPGMFCSTSA-VNSCARCFHFSVCPAGM----- 117
Db 106 VEIRACTKQNRVCAACEAGRYCALKTHSGCRQCMRLSKCGPFGVASSRAPNGVLCKA 165
QY 118 -----IVKPPGTAQKNTVCPSAPGVSPA-----CASPENCK 149
Db 166 CAPCTFSDTTSSTDVCPHRCISILAIPGNASTDAVCAPEPTLSAIPRLVYSQPEPTR 225
QY 150 EPSGGTIPOAKPTVPSPATSSAST-----MPVRGGTGL----- 182
Db 226 SQPLDQEPGPSQTP-SILTSLSGSTPIEOSTKGGISLPIGLIVGVTSLGLMLGLVNCII 284
QY 183 -----AQEAAS-----KLTRAPDPSVSGVRPSDPLSPTQ 213
Db 285 LVORKKKPSCLQRDAKVPVHVPDEKSDQAVGLEQQHLLTAPSSSSSSLESSASAG--DRR 342
QY 214 PCPESGDCRQCEPDYLLDEAGRCTACVSCSRDOLLVEKTPCAWNSRRTCECRPGM---- 269
Db 343 APPGGHQAARWMAQAQGFQE-----ARASSRIDSSSHSGHGH 380
QY 270 ICATSATNSCARCPYPYIPICAGETVTKPQDMAEKDITFEAPPLGTQPCDNP--TPENGEP 327
Db 381 VNVTCIVNVCSSDHSQCSSQASATVGD-----PDAKPSASPRDEQVP 424

QY 328 ASTS--PTQSLVDSQASKTL-----PIPTSAP 353
Db 425 FSQEBECPQSOS---PCETTETLQSHKPLPLGV 454

RESULT 6
148854
gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: I48854
R:Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.
Mamm. Genome 5: 726-727, 1994
A:Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.
A:Reference number: 148854; MUID:9517848; PMID:7873884
A:Accession: I48854
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-459 <RES>
A:Cross-references: EMBL:X76401; NID:9433830; PIDN:CAA53981.1; PID:g433831
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
F:151-188/Domain: NGF receptor repeat homology <NGF>

Query Match 9.5%; Score 198.5; DB 2; Length 459;
Best Local Similarity 20.5%; Pred. No. 0.00021;
Matches 93; Conservative 39; Mismatches 160; Indels 161; Gaps 16;

QY 18 YDKAVRCCYRCYPMGLFPTQCPORPTDCRKQCEPDYLLDEADRCTACVTCSD--DL 74
Db 31 YDKRAQMCCKACPPGQYVHKFNCTSDTVCADEASMYQTVMNFRCLSCSSCGSTQ 90
QY 75 VEKTPCAWNSRVCECRPGMFCSTSA-VNSCARCFHSCVPAGM----- 117
Db 91 VETRACTQONRVCAACEAGRYCALKTHSGSCRCQMRLSKCGFGFVASSRAPNGVLCRA 150
QY 118 -----IVKFGCTAQNIVCEPASGVSPA-----CASPENCK 149
Db 151 CAPGTFSDTSTDVCRPHRISILAIPEGNASTDVCAPESTLSAIPRTLXVSQPEPTR 210
QY 150 EPSSTGTPQAKPTVPSPATSSA-----MPVRGGTRL----- 182
Db 211 SQPLDQEPGSPQTP-SILTSLGSTPIIISQSHKGLSLPLGLVGTSLGLMLGLVNCFI 269
QY 183 -----AQEAAS-----KLTRAPDPSVSRPSSDPGLSPQTQ 213
Db 270 LVQRKKPKCLQORDAKVPHVPDEKSDQAVGLEQHQHLLTTPAPSSSSLESSASAG--DRR 327
QY 214 PCPEGSGCRKQCEPDYLLDEAGRCTACVSCRDDLVEKTPCAWNSRTECRPGM---- 269
Db 328 APPGGHPQARYMAE-----AQSQEARASSRISDSHSGSHGTH 365
QY 270 ICATSATNSCARCVPIPICAGETVTKPQDMAEKDTTFEAPPLGTOPDCNP--TPENGEAP 327
Db 366 VNVTCVIVNCSSDHSQCSQASATVGD-----PDAKPSASPDEQVP 409

QY 328 ASTS--PTQSLVDSQASKTL-----PIPTSAP 353
Db 410 FSQEBECPQSOS---PYETTETLQSHKPLPLGV 439

RESULT 7
T43251
furin (BC 3.4.21.75) - fall armyworm
N:Alternate names: paired basic amino acid cleaving enzyme; proprotein convertase; serin
C:Species: Spodoptera frugiperda (fall armyworm)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T43251
R:Cieplik, M.; Klenk, H.
submitted to the EMBL Data Library, January 1996
A:Description: Cloning and functional characterization of FURIN from Spodoptera frugiper
A:Reference number: 222368
A:Accession: T43251
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-1299 <CIE>
A:Cross-references: EMBL:Z66888; NID:g1167859; PID:e219690; PIDN:CAA93116.1
A:Experimental source: clone Sfurin 6; ovary
C:Function:
A:Description: responsible for the endoproteolytic processing of proproteins with spe
C:Keywords: hydrolase; serine proteinase

Query Match 9.5%; Score 198.5; DB 2; Length 1299;
Best Local Similarity 21.7%; Pred. No. 0.00051;
Matches 97; Conservative 39; Mismatches 116; Indels 195; Gaps 22;

QY 10 TCHGN-----PSHYDKAVRCCYRCYPMGLFPTQCPORPTDCRKQCEPDY 56
Db 864 TCHPSTYALDGRVCVTSCPPAYADKK-RKECMRCPVG-----CSTCTSAFCLSCPEKWE 916
QY 57 LDE-----ADRCTA-----CVTCSRDDLVEKTPCAW 82
Db 917 LNKKGKMPVSGDKCSAGFAVDQKRCRNPACDSCYGENEGHCLTCPNLLQDYKC-- 974
QY 83 NSSRVCECRPGMFCSTSAVNSCARCFHSCVPAGMIVKFPCTAQNKTVCPEPASGVSPAC 142
Db 975 -----VPECSKGY--AEAGRCARCMH-----GCSD-C 999
QY 143 ASPENCKEPPSGTIPQAKPTVPSPATSSASTMPVRGGTRLAQEAASKLTRAPDPSVSG- 201
Db 1000 VSLNC-----TSCATLRLQSG-----ACRTSCADGYADRGT 1033
QY 202 -----RPSSDGLSPTOPCEG-----SGDCR----- 223
Db 1034 CSKYLSRCTICIGRRDOQASCPEGWRLAAGECHPECPQGFYQSPGCRHCHHYCRECDG 1093
QY 224 -----KQCEPDYLLDEAGRCTACVSCRDDLVEKTPCAWNS--RTCECRPGMICATSAT 276
Db 1094 SGPLHCKSCPPRMFLD--GLLCMECLGSQYYDATSGTCRSCDASCRTCS--GFGQFSCT--- 1148
QY 277 NSCAR-----CVPIPICAGETVTKPQDMAEKDTTFEAPPLGTOPDCNPTPENGEA 326
Db 1149 -TCSRPLRLDRLNNQCVP---CC-----SERGVNTNSTPTDC---CHCNPENGECEC 1191
QY 327 PASTSTQSLVDSQASKTLPIPTSAP 353
Db 1192 INSSVAGKRRIAEWGAHTAPSAADAAP 1218

RESULT 8
JC4302
tumor necrosis factor receptor p55 precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 29-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 23-Jul-1999
C:Accession: JC4302; PC4093
R:Suter, B.; Pauli, U.
Gene 163, 263-266, 1995
A:Title: Cloning of the cDNA encoding the porcine p55 tumor necrosis factor receptor.
A:Reference number: JC4302; MUID:96011645; PMID:7590278
A:Accession: JC4302
A:Molecule type: mRNA
A:Residues: 1-461 <SUT>
A:Cross-references: GB:U19994; NID:g1141752; PIDN:AAC48499.1; PID:g1141753
A:Accession: PC4093
A:Molecule type: protein
A:Residues: 1-7 <SU2>
A:Experimental source: kidney cell line 15
C:Genetics:
A:Gene: tnfr
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
C:Keywords: glycoprotein; kidney; receptor; transmembrane protein; tumor
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-461/Product: tumor necrosis factor receptor p55 #status predicted <MAT>
F:44-194/Domain: extracellular cysteine rich #status predicted <EXT>
F:44-82/Domain: NGF receptor repeat homology <NGL>
F:84-126/Domain: NGF receptor repeat homology <NGF>
F:211-231/Domain: transmembrane #status predicted <TM>

Db 4120 QCRDQCPGPTCKES-NDSSRTCAKIPL 4147

RESULT 11

A55677
laminin beta-2 chain precursor (version 1) - human.
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 17-Mar-1999
C:Accession: A55677
R:Wewer, U.M.; Gerecke, D.R.; Durkin, M.E.; Kurtz, K.S.; Mattei, M.G.; Champliand, M.F.;
Genomics 24, 243-252, 1994
A:Title: Human beta2 chain of laminin (formerly S chain): cDNA cloning, chromosomal loca
A:Reference number: A55677; MUID:95213013; PMID:7698745
A:Accession: A55677
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1797 <NEW>
A:Cross-references: GB:X79683
C:Genetics:
A:Gene: GDB:LAMB2
A:Cross-references: GDB:132363; OMIM:150325
A:Map position: 3p21.3-3p21.2
C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
C:Keywords: basement membrane; extracellular matrix; glycoprotein; heterotrimer
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-1797/Product: laminin beta-2 chain #status predicted <MAT>
F:283-344/Domain: laminin-type EGF-like homology <LE01>
F:347-407/Domain: laminin-type EGF-like homology <LE02>
F:410-467/Domain: laminin-type EGF-like homology <LE03>
F:470-519/Domain: laminin-type EGF-like homology <LE04>
F:522-552/Domain: laminin-type EGF-like homology #status atypical <LE05>
F:783-828/Domain: laminin-type EGF-like homology <LE06>
F:831-874/Domain: laminin-type EGF-like homology <LE07>
F:877-924/Domain: laminin-type EGF-like homology <LE08>
F:927-982/Domain: laminin-type EGF-like homology <LE09>
F:985-1034/Domain: laminin-type EGF-like homology <LE10>
F:1037-1091/Domain: laminin-type EGF-like homology <LE11>
F:1094-1139/Domain: laminin-type EGF-like homology <LE12>
F:1142-1186/Domain: laminin-type EGF-like homology <LE13>

Query Match 8.2%; Score 171; DB 2; Length 1797;

Best Local Similarity 23.2%; Pred. No. 0.023;
Matches 89; Conservative 30; Mismatches 145; Indels 120; Gaps 19;

QY 1 FPQDRPFEDTCHGNPSHYVDKAVRRC-----CYRCPMGLFPTQCPQ-----RPTDCR 48
Db 871 FPCRPVCNCHADCNHTGACLCRDLTGECERCIAFGHDPRLPYGAQCRPCPCP 930
QY 49 K-----QCEPDYYLDE-----ADRCTACTCSDLDLVEKTPCAWSSRYC 88
Db 931 EPGSGQRHFATSCHQDEYSQQIVCHCRAGYTGRCACAPGQFGD-----PSRPGACQLC 985
QY 89 EC-----RGCMFCSTSAVNSCARCPFHVSVCPP--AGMIVFPPTAQNTV-- 130
Db 986 ECGNIDPMDPADCPHPG-----QCLRLHHTEGPHCAHSGKPGFHGOAARQSCHR 1036
QY 131 CEPASPGVSP-ACASPENCK-EPSSSTIP-----QAKPTVPSPATSSASTMPVRGTRLA 183
Db 1037 CTCNLGNTPOQCPSPDQCHDQPSGQCPCLPNVOA-----KQCEPD 229
QY 184 QEAAS-----KLTRAPDSSVSGVRSPDGLSPQPCPEGSGDCR-----KQCEPD 229
Db 1075 VDRCAPNFWNLTSGHGQCPACLPREBEG--PT--CNFTGCHPGAGFGGRTCEQEL 1130
QY 230 YYLDEAGRCTACVSGSRDDLVEKTPCAWSSRTCECRPGMI-----CATSATNSCARCV 283
Db 1131 HWGDPGLQCHACDCDSRG---IDTPQCHRTGHCCTCRPGVSGVRCDQCARGFSGIFPACH 1187
QY 284 PYPICAGETVTKPQDMAEKDTFFE 307
Db 1188 PCHACFGDMRWVQDLAARTORLE 1211

RESULT 12

A55494
latent transforming growth factor-beta-binding protein - human
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 11-Jan-2000
C:Accession: A55494
R:Moren, A.; Olofsson, A.; Stenman, G.; Sahlin, P.; Kanzaki, T.; Claesson-Welsh, L.;
J. Biol. Chem. 269, 32469-32478, 1994
A:Title: Identification and characterization of LTRBP-2, a novel latent transforming g
A:Reference number: A55494; MUID:95096101; PMID:7798248
A:Accession: A55494
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1820 <MOR>
A:Cross-references: GB:X37976
C:Genetics:
A:Gene: GDB:LTRBP2
A:Cross-references: GDB:568901
A:Map position: 11pter-11qter
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:1222-1257/Domain: EGF homology <EGF>
F:1525-1565/Domain: EGF homology <EGF1>

Query Match 8.0%; Score 166; DB 2; Length 1820;

Best Local Similarity 22.0%; Pred. No. 0.044;

Matches 89; Conservative 33; Mismatches 152; Indels 130; Gaps 20;

QY 40 CQORPT-----DCRKOCEPDYYLDEADRCTACTVCTCSRDDLVKTPCA-WNSSRVCECRPG 93
Db 592 CPPRPASPVIENGQLECPQGYKRLNLTHTCODINECLTLGLCKDAECVNTRGSLYCTCRPG 651
QY 94 M-----FCSTSASV-----NSCAR 106
Db 652 LMLDPSRSCVSKAISMLOGLCYRSILGPGCTCTPLPLAQRTKQICCCSRVKGWGSCEK 711
QY 107 C-----FFHVSVCAGMIVKFPCT-----AKNTVCBPASPGVSPACASPNCKEPS 152
Db 712 CPLPGTEAFREICPAGHYTYASSDIRLSMRKAEELARP-----PREQOGRS 760
QY 153 SGTIP-QAKPTVPSPATSS---ASTMPVRGTRLAQEAASKLTRAP----- 194
Db 761 SGALPGPAERQPLRVVDTWLEAGTIPDKGDSQAGQVTTT-VTHAPAWVTGNATPPMPE 819
QY 195 ----DSPSSVSGRPSDD-----PGL-----SPTQPCPEGSGDC-----RKOCEPDY 231
Db 820 QGIAETQEQVTPSTDVLYTLSTPGIDRCAGATNVC--GPGTCVNLPGYRCVCSFGYQ 877
QY 232 LDEA-GRCTACVSCSRDDLVEKTPCAWSSRTCECRPGMICATS-ATNSCA---RCVPY 285
Db 878 LHPSQAYCTDDNECLRDCQKGRGINRVGYSYSCFCYVGYTLATSGATQECQDINECEQP 937
QY 286 PICAGETVTKPQDM--AEKDTTFEAPPLGTQPCNTPPENGEP 327
Db 938 GYCSGGQCTTEGSHYCECDQGYIMVRKHCQDINECRHFGTCGP 981

RESULT 13

T42017
cysteine rich protein - Giardia intestinalis
C:Species: Giardia intestinalis
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T42017
R:Chen, N.; Upcroft, P.; Upcroft, J.
Parasitology 111, 423-431, 1995
A:Title: A Giardia duodenalis gene encoding a protein with multiple repeats of a toxi
A:Reference number: 222027
A:Accession: T42017
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1274 <CHE>
A:Cross-references: EMBL:L29079; NID:g951190; PID:g951191; PIDN:AAA74587.1
A:Experimental source: specific host: Homo sapiens


```

Db 1042 LL-----GTNPQQCPSPDQCHCDPSSGQCP-CLPNVQGPSC 1076
QY 169 SSASTMPVRGGTRLAQAASKLTRAPDPSVSVGRPSSDGLSPTQPCPEGSGDCR----- 223
Db 1077 D-----RCAPNFWN-LTSGHGQCPACHPFSRARG--PT--CNEFTGQCHCRAGF 1120
QY 224 -----KOCEPDYLLDEAGRCTACVSCSRDDLVEKTPCAWNSRSTCECRPGMI-----CA 272
Db 1121 GGRTCEQELHWGDPGLQCHACDCDSRG---IDTPQCHRTGHCSCRPVSGVRCDDQCA 1177
QY 273 TSATNSCARCVYPICAGETVTKPQDMAEKDTTFE 307
Db 1178 RAFSGIFFACHPCHACFGDWDVRVODLAARTQRL 1212

```

Search completed: December 11, 2002, 15:24:51
 Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:19:26 ; Search time 35 Seconds
(without alignments)
2189.986 Million cell updates/sec

Title: US-09-628-126-2_COPY_19_390

Perfect score: 2080

Sequence: 1 FPQDRPFEDTCHGNPSHYD.....PVALSSSTGKPVLDAGPVLFV 372

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_bacheap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	198.5	9.5	459	11 062327	Q62327 mus musculus
2	198.5	9.5	1299	5 Q26489	Q26489 spodoptera
3	197	9.5	433	11 0912M6	Q912m6 rattus norv
4	193.5	9.3	482	11 088734	Q88734 mus musculus
5	193	9.3	2284	5 Q3VPGL1	Q9vpgl1 drosophila
6	180.5	8.7	111	12 P87599	P87599 cowpox viru
7	177.5	8.5	739	5 Q9GS24	Q9gs24 giardia lam
8	176	8.5	3767	5 Q3UAI3	Q9ual3 caenorhabdi
9	175.5	8.4	110	12 Q8QN46	Q8qn46 cowpox viru
10	174.5	8.4	348	12 057103	057103 monkeypox v
11	174.5	8.4	348	12 057108	057108 monkeypox v
12	174.5	8.4	348	12 057277	057277 monkeypox v
13	174.5	8.4	350	12 057123	057123 cowpox viru
14	172	8.3	349	12 057100	057100 monkeypox v
15	171.5	8.2	351	12 057117	057117 cowpox viru
16	170	8.2	349	12 057102	057102 monkeypox v

17	170	8.2	349	12 057291	057291 monkeypox v
18	169.5	8.1	1764	11 039806	039806 rattus norv
19	169.5	8.1	5374	11 099ND0	099nd0 mus musculus
20	168.5	8.1	765	5 Q9NL50	Q9nl50 sarcophaga
21	168	8.1	349	12 057101	057101 monkeypox v
22	168	8.1	946	10 022015	022015 cylindrothe
23	166.5	8.0	856	12 080UT7	080ut7 infectious
24	- 166	8.0	349	12 057099	057099 monkeypox v
25	166	8.0	1821	4 Q14767	Q14767 homo sapien
26	165.5	8.0	1154	5 Q9GQ46	Q9gq46 giardia lam
27	165.5	8.0	1274	5 Q9NGL3	Q9ngl3 giardia lam
28	165.5	8.0	1274	5 Q24977	Q24977 giardia lam
29	165.5	8.0	1428	5 044341	044341 haliotis ru
30	165	7.9	283	6 Q9XS28	Q9xs28 cercopithec
31	165	7.9	807	5 018511	018511 trichoplusi
32	165	7.9	1833	11 008999	008999 mus musculus
33	164.5	7.9	351	12 073559	073559 cowpox viru
34	164.5	7.9	355	12 085308	085308 cowpox viru
35	164	7.9	969	4 Q9NS51	Q9ns51 homo sapien
36	163.5	7.9	326	12 057122	057122 cowpox viru
37	163	7.8	969	4 Q96KG6	Q96kg6 homo sapien
38	162	7.8	186	12 072735	072735 cowpox viru
39	161.5	7.8	572	5 Q19594	Q19594 caenorhabdi
40	161.5	7.8	949	5 P90956	P90956 caenorhabdi
41	160.5	7.7	326	12 057120	057120 cowpox viru
42	160.5	7.7	5146	6 Q8SPM4	Q8spm4 bos taurus
43	160	7.7	1963	6 Q28019	Q28019 bos taurus
44	159.5	7.7	351	12 057121	057121 cowpox viru
45	159.5	7.7	542	2 068872	068872 myxococcus

ALIGNMENTS

RESULT 1

ID	Q62327	PRELIMINARY;	PRT;	459 AA.
AC	Q62327;			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)		
DE	Murine tumour necrosis factor receptor 2 protein (Fragment).			
GN	TNFRSF1B.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NOD;			
RA	Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;			
RT	"Amino acid variation in the tumor Necrosis factor receptor 2 is linked to autoimmune diabetes in NOD mice.";			
RL	Genomics 0:0-0(0).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NOD;			
RX	MEDLINE=95178848; PubMed=7873884;			
RA	Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;			
RT	"Allelic variation of the type 2 tumor necrosis factor receptor gene.";			
RL	Mamm. Genome 5:726-727(1994).			
DR	EMBL; X76401; CAA53981.1; -.			
DR	HSSP; P19438; INCF.			
DR	MGI; MGI:1314883; Tnfrsf1b.			
DR	InterPro; IPR001368; TNFR_c6.			
DR	Pfam; PF00020; TNFR_c6; 4.			
DR	SMART; SM00208; TNFR; 4.			
DR	PROSITE; PS00652; TNFR_NGFR_1; 2.			
DR	PROSITE; PS50050; TNFR_NGFR_2; 3.			
KW	Receptor.			
FT	NON_TER	1	87	S -> T.
FT	VARIANT	87	87	

FT VARIANT 93 268 T -> I.
FT VARIANT 268 345 F -> I.
FT VARIANT 345 421 S -> F.
FT VARIANT 421 459 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CFC1C4626DF CRC64;

Query Match 9.5%; Score 198.5; DB 11; Length 459;
Best Local Similarity 20.5%; Pred. No. 1.2e-07;
Matches 93; Conservative 39; Mismatches 160; Indels 161; Gaps 16;

QY 18 YDKAVRCYRCYCPMGLFTQOCQPCQPTDCRKCQCPDYVLDLDEADCTACTVTCSDR---DL 74
DB 31 YDKAQMCAKCPGQYVHFCNKTSDTVACDCASMTQVWNGFRTCLSCSSCSTDQ 90
QY 75 VEKTPCAMNSRVRCRPFMGFSTSA-VNSCARCFHVSVCYPAGM-----117
DB 91 VETRACTQONRVCAACEAGRYCALKTHSGSCRCQRLSKCGPGFGVASSRAPNGVNLCKA 150
QY 118 -----IVKFGTAQKNVCEPASGVSPA-----CASPENCK 149
DB 151 CAPGTFSDTSSDVCPRHRIGSILAIPEGNASTDVACAPESPTLSAIPRTLYVSQPEPTR 210
QY 150 EPSSTIQAQKTPVSPATSSAST---MPVRGGTRL-----182
DB 211 SQPLDQEPGSPQTP-SILTSLSGSTPIIQSTKGGISLPIGLIVGVTSLGLMLGLVNCFI 269
QY 183 -----AQEAAS-----KLTRAPDSPSSVGRPPSDPGLSPQ 213
DB 270 LVQRKKKPSCLQDAKVHPVDEKSDQAVGLEQHLITAPSSSSSSLESSASAG--DRR 327
QY 214 PQEBSGCRKOCEPDYLDLDEAGRTACTVCSRRDDLVEKTPCAMNSTRTCERPGM-----269
DB 328 APPGGHPQARVMAE-----AQSQEARASSRISDSHSGHGH 365
QY 270 ICATSATNSCARCVYPICAGETVTKPQDMAEKDTTFEAPPLGTQPCDNP--TPENGEP 327
DB 366 VNVTCIVNVSSSHSSQSOASATVGD-----PDAKPSASPKDEQVP 409
QY 328 ASTS--PTOSLLVDQASKTL-----PIPTSAP 353
DB 410 FSQEECPQSQSQS---PYETETLQSHKPLPLGVP 439

RESULT 2

Q26489 PRELIMINARY; PRT; 1299 AA.
AC Q26489;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Endoprotease furin.
GN FURIN.
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=7108;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SF9;
RA Cieplik M., Klenk H.;
RT "Cloning and functional characterization of FURIN from Spodoptera frugiperda (sf9) cells."
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 268888; CAA93116.1; -;
DR HSSP; Q99405; IMPT.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR002884; P_domain.
DR Pfam; PF01483; P; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR ProDom; PD000717; P_domain; 1.

DR SMART; SM00261; FU; 10.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease.
SQ SEQUENCE 1299 AA; 142020 MW; 4C3799C7BBC572AB CRC64;

Query Match 9.5%; Score 198.5; DB 5; Length 1299;
Best Local Similarity 21.7%; Pred. No. 3.5e-07;
Matches 97; Conservative 39; Mismatches 116; Indels 195; Gaps 22;

QY 10 TCHGN-----PSHYVDKAVRRCYRCYCPMGLFTQOCQPCQPTDCRKCQCPDY 56
DB 864 TCHPSTYALDGRVTSCTPPAYADKK-RKECMRCPVG-----CSTCTSAFCLSCPEKWE 916
QY 57 LDE-----ADRCTA-----CVTCSRDDLVEKTPCAW 82
DB 917 LNKKGKMPVGSCKSAGEFAVDQCKRCNPACDSCYGENEGHCLTCPNPDLQDYKC-- 974
QY 83 NSRVCECRPGMFCSTSAVNSCARCFHVSVCYPAGMIVKFGTAQKNVCEBPASGVSPAC 142
DB 975 ---VPECSKGY--AEAGRCARCMH-----GCS-D-C 999
QY 143 ASPENCKEPPSGTIPQAKPTVPSPATSSASTMPVRGTRLAQEAASKLTRAPDSPSVG- 201
DB 1000 VSLNC-----TSCASTLRQSG-----ACRTSCADGYADRG 1033
QY 202 ---RPSSDPGLSPTQCPPEG-----SGDCR-----223
DB 1034 CSKCYLSRCTICPRDQACSCPEGRWLAAGECHPECPQGYSPGCGRCHHHCYCRECDG 1093
QY 224 -----KQCEPDYLDLDEAGRTACTVCSRRDDLVEKTPCAMNNS-RTCECRPGMICATSAT 276
DB 1094 SGPLHCKSCPPREMLD-GGLCMECLGQYVYDATSGTCRSCDASCTCS-GPGQFSCT---1148
QY 277 NSCAR-----CVYPICAGETVTKPQDMAEKDTTFEAPPLGTQPCDNPENGEA 326
DB 1149 -TCSRPLRIDRLNQCQVP---CC-----SERGVNTSTPTDTC--CHCNPENGEC 1191
QY 327 PASTSPOTOSLLVDQASKTLPIPTSAP 353
DB 1192 INSSVAGKRRIAEWGAHTAPSADAAP 1218

RESULT 3

Q912M6 PRELIMINARY; PRT; 433 AA.
ID Q912M6
AC Q912M6;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Tumor necrosis factor receptor type II (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Osburg B., Peiser C., Doemling D., Schomburg L., Voigt K., Bickel U.;
RT "TNF-receptors p60 and p80 are constitutively expressed by rat brain capillary endothelial cells and participate in TNF-alpha transport through the blood-brain barrier."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF420214; AAL16021.1; -;
DR InterPro; IPR001368; TNFR_C6.
DR Pfam; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 433 433
SQ SEQUENCE 433 AA; 45723 MW; 75736D835E72CA4A CRC64;

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulo G., Milchina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003591; AAF51592.1; -
 DR HSP: P01180; 2BN2
 DR FlyBase: FBgn0036985; CG5847.
 DR InterPro: IPR001507; Endoglin/CD105.
 DR SMART: SM00241; ZP; 1.
 SQ SEQUENCE 2284 AA; 248295 MW; F495BA8EC8A4674A CRC64;

Query Match 9.3%; Score 193; DB 5; Length 2284;
 Best Local Similarity 22.7%; Pred. No. 1.7e-06;
 Matches 93; Conservative 49; Mismatches 160; Indels 108; Gaps 23;

QY 27 CYRCPMGLFPTQOCQORPTDCR-----KQCEPDYVLDADRCT-----ACVTCS 70
 DB 898 CLUNCYFG-SPDPRCKVPTTKAGGFDGSDPKQCPATYLPSTRTPTIAKPCRYCGS 956
 QY 71 RD-----DLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHSCVCPAGMIVKFFGTAA 126
 DB 957 KDPGCPQITRAPTSRPR--CYPG-----STDPEC-----HPTSSPAITRIPVTR 1003
 QY 127 -----KNTVCEPAS--PGVSPACAPENCKEFS--SGTI-POAKPTPVSPATSSASTMPV 176
 DB 1004 IPLTTAKPCRYCGSEPGCQAPTRPITTSKPCRYGSLDPECQSTYLPPTPVTRTPT 1063
 QY 177 RGGTLAQAASAKLRAPDSSVGRSDPGLSTQPCP-EGSGDCRKCQCPDYVLEDA 235
 DB 1064 IPTTRI-----PVTTSKPCRYG-SDTRCRKPEPVTPKPCRYGSPNDEC 1108
 QY 236 GRCTACVSCSDDLVEKTPCAW--NSSRTC-----ECRPGMICSATNSCARCV 283
 DB 1109 QKATYSPPTTRPTVTTSKPCRYGSDSRCPQKPTTLKPKCSG-----SSDPECLN 1163
 QY 284 ---PYPICAGETVTK-----PQDMAEKDTTF-----EAPPLGTQPCDNTPEGEAP 327
 DB 1164 PGSPDPRCPKVPTTKKSGCFDGSQDPRCPQATYLPSPSSRRPPTTAPKPCRYGSDPSCP 1223
 QY 328 ASTSP-TQSLLDVDSQASK-----TLPIPTSA-PVALSSTGKP 362
 DB 1224 QPTQPIITRLPITTKARCYGCPGSTAPEQAPATPTPTSTRTPTTTTTSKP 1273

RESULT 6
 P87599
 ID P87599 PRELIMINARY; PRT; 111 AA.
 AC P87599;
 DT 01-MAY-1997 (TremBLrel. 03, Created)
 DT 01-MAY-1997 (TremBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
 DE 41KB fragment from LEFT end of genome.
 GN C5L OR D13L.
 OS Cowpox virus (CPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10243;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GRI-90;
 RX MEDLINE=97068532; PubMed=8963248;

RA Sazonov P.F., Petrov N.A., Riazankina O.I., Totmenin A.V.,
 RA Shchelkunov S.N., Sandakhchiev L.S.:
 RT "Genes of a circle of hosts for the cowpox virus."
 RN Dokl. Akad. Nauk 349:829-833(1996).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=GRI-90;
 RX MEDLINE=98229462; PubMed=9568042;
 RA Shchelkunov S.N., Sazonov P.F., Totmenin A.V., Petrov N.A.,
 RA Riazankina O.I., Gutorov V.V., Kotwal G.J.:
 RT "The genomic sequence analysis of the left and right species-specific
 RT terminal region of a cowpox virus strain reveals unique sequences and
 RT a cluster of intact ORFs for immunomodulatory and host range
 RT proteins."
 RL Virology 243:432-460(1998).
 DR EMBL: X94355; CAA64084.1; -
 DR EMBL: Y11842; CAA72589.1; -
 DR InterPro: IPR001368; TNFR_C6.
 DR Pfam: PF00020; TNFR_C6; 1.
 DR SMART: SM00208; TNFR; 1.
 DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
 SQ SEQUENCE 111 AA; 12162 MW; 771D9A59D671CF21 CRC64;

Query Match 8.7%; Score 180.5; DB 12; Length 111;
 Best Local Similarity 39.6%; Pred. No. 6.8e-07;
 Matches 42; Conservative 9; Mismatches 34; Indels 21; Gaps 4;

QY 45 TDCRKOCEPDYVLDADRCTACVTCRDDLVKTPCAWNSRVCECRPGMFCSTSAVNS 103
 DB 18 TSEGTCTPNYLYLEPDGICTACVTC-LSNMVETIQCGDPKPKKCGGLKCTVPAVNS 76
 QY 104 CARCFHSCVCPAGMIVKFFGTAKNTVCEPASPGVSPACASPENCK 149
 DB 77 CARC-----TPDTTIKIEPTDQ-----CCTTPDNTK 103

RESULT 7
 Q9GS24
 ID Q9GS24 PRELIMINARY; PRT; 739 AA.
 AC Q9GS24;
 DT 01-MAR-2001 (TremBLrel. 16, Created)
 DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE Variant-specific surface protein vsp9p10.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WB;
 RX MEDLINE=21117019; PubMed=11179375;
 RA Nash T.E., Lujan H.T., Mowatt M.R., Conrad J.T.:
 RT "Variant-Specific Surface Protein Switching in *Giardia lamblia*."
 RL Infect. Immun. 69:1922-1923(2001).
 DR EMBL: AY007596; AAG16629.1; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR005127; Giardia_VSP.
 DR Pfam: PF03302; VSP; 2.
 DR SMART: SM00261; FU; 8.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 SQ SEQUENCE 739 AA; 76763 MW; 511A78644EBCEC15 CRC64;

Query Match 8.5%; Score 177.5; DB 5; Length 739;
 Best Local Similarity 21.2%; Pred. No. 8.2e-06;
 Matches 89; Conservative 42; Mismatches 123; Indels 165; Gaps 24;
 QY 27 CYRCPMGLFPTQOCQORPTDCRQ-----GEPDYVLEDA-DRCTACVTCRDDLVKTP 79
 DB 164 CTRC-----NAPTEADANAKAATCTACQADRYLKTAQDQATSCVTEQECTGTGEFF 215
 QY 80 CAWNSRVCE-----CRPGMFCSTSAV--NSC-----ARCFHSCVCPAGM 117
 RX MEDLINE=97068532; PubMed=8963248;

```
Db 216 AKNGTKTCEACETCKT---CKTEAAKCTCKDDKPYLLKONGSDTGTCTVTEAECKTGN 272
QY 118 IVKFPQTAGNTVCEPASPGVPACASPENCKPSSGTP-----QAKPTPV----- 165
Db 273 -THYPDTEPKT-CKTASGTSAGCAT---CEKAGDGAACKTCGKQKKVQPNKKGCVEN 327
QY 166 -PATSSASTMPVRGTF-----RLAQEAASKLTLRAPDSPSSVGRPSDDPG----- 208
Db 328 CPSSNSDKKTP---GTCECEGVYVPEAGTGTCKKPPDPA---PCNTPGCKTCTSEPKTS 381
QY 209 -----LSPQPC-----PESSGDCRQCEPDYLL--DEAGRCTA 240
Db 382 KEVCTCEDPKALTGQCIYCEHLEGYYEGTSEGGKACKKEVEENCLLNCNOCET 441
QY 241 C-----VSCSRDLVEKTPCAWNSRTC-----ECRPGMI 270
Db 442 CKDGYKKTACAKDTSCKT-CANGPNCTSCAEKAKKALSVGEENGTGCKSECKPG-- 498
QY 271 CATSAINSCARCPVYPICAGETVTRKPDMAEKDITTFEAPPLGTQPDNPTPENGAPAS 329
Db 499 -----TNNCKC-----ELTVDTGTAY---CSKCKDANQFPQNGVCSAA 533

RESULT 8
Q9UAI3 PRELIMINARY; PRT; 3767 AA.
AC Q9UAI3; Q21340;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Transmembrane cell adhesion receptor MUA-3 precursor (K08E5.3 protein)
DE (Fragment).
GN MUA-3 OR K08E5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RA Lu Z., Vogel B., Hedgecock E.;
RT "mua-3 mRNA Splicing Pattern Revealed.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Kershaw J.K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
[4]
RP SEQUENCE FROM N.A.
RA Berks M., Smith A., Lloyd C.R.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF139060; AAD29428.1; -.
DR EMBL; Z30974; CAA83226.2; -.
DR EMBL; Z30423; CAA83226.2; JOINED.
DR EMBL; Z30423; CAC42345.1; -.
DR EMBL; Z30974; CAC42345.1; JOINED.
DR HSSP; P01130; 1L DL.
DR InterPro; IPR001152; Asx_hydroxyl.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR000082; SEA_domain.
DR InterPro; IPR002035; VWF_A.
Pfam; PF00008; EGF; 33.
DR Pfam; PF00057; ldl_recept_a; 3.
```

```
DR Pfam; PF01390; SEA; 2.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO0289; DISINTEGRIN.
DR PRINTS; PRO0261; LDLRECEPTOR.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00001; EGF_like; 45.
DR SMART; SM00192; LDLA; 4.
DR SMART; SM00200; SEA; 2.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_32.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 6.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00068; LDLRA_2; 2.
DR PROSITE; PS00024; SEA; 4.
DR PROSITE; PS0234; VWFA; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Receptor; Repeat;
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 >3767 TRANSMEMBRANE CELL ADHESION RECEPTOR MUA-3.
FT NON_TER 3767 3767
SQ SEQUENCE 3767 AA; 417284 MW; 8DA3AE5EA50AEB8E CRC64;

Query Match 8.5%; Score 176; DB 5; Length 3767;
Best Local Similarity 22.2%; Pred. No. 5.8e-05;
Matches 91; Conservative 44; Mismatches 162; Indels 112; Gaps 21;

QY 6 PFEDTCHGNPSHYDYKAVRRCCYRCPMG---LFTTQ---CPQPTDCRQCEPDYLLD 58
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 868 PRLNTC-GKNAECIDLAEGYTC-QCRSGYADISPVSPGRICARVNECSN--KEYNVND 923

QY 59 EADRCTACVTCSDRLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHSHVCPAGMI 118
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 924 -----CSENAICADTEHSYS---CRCRPGFADVSAFN----- 953

QY 119 VKFPGTAQNTVCEPASPGVPACASPENCKPSSGTPQAKPTPVSPATSSATMPVRG 178
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 954 -KLPGRCIEAVNECASPSLND-CSKNAFCEDAKEGICTCRPGYVDSNPAA-----RH 1006

QY 179 GTRLAQEAASKL-TRAPDSPSSVGRPSDGLSTQP-----CPESGDCRQCEPDY 230
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1007 PGRICIKPVEKIKTDLKDTSF-----STDGCDPKPKCGANEACVQPHGOHCECVETA 1061

QY 231 YLDEAGRCTACVSCSRDLDVEKTPCAWN--SSRTCECRPGMI-----CATSAT 276
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1062 FRYTDGSCRYSACSKRNTCDKNAICLNRFDSYTCQCRPGYIDLSDLTNAPGRICKELI 1121

QY 277 NSCA----RCVPYPICA---GETVTKPDMAEKDITTFEAPP-----LGT 313
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1122 NECASSDNECSYARCIDATNGACQCLDGFIDYSSRYNKPGRQCTNSNNECKSEKSLNT 1181

QY 314 ---QPCNPTPENGAPASTPTQSLLVDSQASKTLP-----IPTSAP 353
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1182 CDENADCVDTPD-----GYTCQCYGGFVDVSSNANLPPGRVCTVQTTCP 1225

RESULT 9
Q8QN46 PRELIMINARY; PRT; 110 AA.
AC Q8QN46;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE V014.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIGHTON RED;
```



```

RX MEDLINE=83117629; PubMed=6961398;
RA Pickup D.J.; Bastia D., Stone H.O., Joklik W.K.;
RT "Sequence of terminal regions of cowpox virus DNA: arrangement of
RT repeated and unique sequence elements.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:7112-7116(1982).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIGHTON RED;
RA MEDLINE=90177240; PubMed=2309453;
RX Parsons B.L., Pickup D.J.;
RT "Transcription of orthopoxvirus telomeres at late times during
RT infection.";
RL Virology 175:69-80(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIGHTON RED;
RX MEDLINE=91196263; PubMed=2014645;
RA Hu F.Q., Pickup D.J.;
RT "Transcription of the terminal loop region of vaccinia virus DNA is
RT initiated from the telomere sequences directing DNA resolution.";
RL Virology 181:716-720(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIGHTON RED;
RX MEDLINE=94378510; PubMed=8091665;
RA Hu F.Q., Smith C.A., Pickup D.J.;
RT "Cowpox virus contains two copies of an early gene encoding a soluble
RT secreted form of the type II TNF receptor.";
RL Virology 204:343-356(1994).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIGHTON RED;
RA Pickup D.J.;
RT Submitted (AUG-1993) to the EMBL/GenBank/DBDJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIGHTON RED;
RA Dietrich F.S., Ray C.A., Sharma A.D., Allen A., Pickup D.J.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBDJ databases.
DR EMBL: AF482758; AAM13462.1; -
SQ SEQUENCE 110 AA; 12040 MW; 5F61166C7A43ACEC CRC64;

Query Match 8.4%; Score 175.5; DB 12; Length 110;
Best Local Similarity 38.6%; Pred. No. 1.6e-06;
Matches 39; Conservative 11; Mismatches 28; Indels 23; Gaps 5;

QY 224 KQCEPDYILD-EAGRCTACVSCSRDLVEKTPCAWNSRTCECRGMCATSAFNSCARC 282
Db 22 KTCPADYYLEPEDGLCTACVTC-LSNMVEVETQSCGPKRKQCQGLKCTVPAVNSCARC 80
QY 283 VPYPICAGETVTKPDMAEKDTFEAPLGTQPCDNPEN 323
Db 81 TP-----DITTK---KVQKE-----QCCTPDN 100

RESULT 10
057103
ID 057103 PRELIMINARY; PRT; 348 AA.
AC 057103;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE Tumor necrosis factor receptor II homolog.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAIRE-1970;
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RT "DNA sequence analysis as a criterion for allocation of the
RT orthopoxviruses to a particular species.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBDJ databases.
DR EMBL: U88142; AAB94367.1; -
DR HSP: O14763; IDOG.
DR InterPro: IPR001368; TNFR_c6.
DR Pfam: PF00020; TNFR_c6; 2.
DR SMART: SM00208; TNFR; 2.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW Receptor.
SQ SEQUENCE 348 AA; 38212 MW; E555979057DEC91F CRC64;

Query Match 8.4%; Score 174.5; DB 12; Length 348;
Best Local Similarity 26.9%; Pred. No. 6.4e-06;
Matches 36; Conservative 22; Mismatches 71; Indels 5; Gaps 3;

QY 26 CCYRCPMGLFPTQCCPQRPTDCRKQCEPDYLLDEADRCTACVTC---RDDLVEKTPCAW 82
Db 43 CCLSCPPGPGTYASRLCDSKTNTQCTPCGSDTFTSHNNHLQACLSCNGRCDNQVETRSCNT 102
QY 83 NSSRVCECRPGMFCSTSAVNSCARCFHVSVCAPAGMLVKFPGTAQKNVCEPASPGV-SPA 141
Db 103 THNRICECSPGYYCLLKSGSGCRTCISKTKCGIGYGS-GYTSTGDVICSPCGPGTYSHT 161
QY 142 CASPENCKEPSSGT 155
Db 162 VSSTDKCEPVTSTNT 175

RESULT 11
057108
ID 057108 PRELIMINARY; PRT; 348 AA.
AC 057108;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE Tumor necrosis factor receptor II homolog.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAIRE-1970;
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RT "DNA sequence analysis as a criterion for allocation of the
RT orthopoxviruses to a particular species.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBDJ databases.
DR EMBL: U88142; AAB94367.1; -
DR HSP: O14763; IDOG.
DR InterPro: IPR001368; TNFR_c6.
DR Pfam: PF00020; TNFR_c6; 2.
DR SMART: SM00208; TNFR; 2.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW Receptor.
SQ SEQUENCE 348 AA; 38212 MW; E555979057DEC91F CRC64;

Query Match 8.4%; Score 174.5; DB 12; Length 348;
Best Local Similarity 26.9%; Pred. No. 6.4e-06;
Matches 36; Conservative 22; Mismatches 71; Indels 5; Gaps 3;

QY 26 CCYRCPMGLFPTQCCPQRPTDCRKQCEPDYLLDEADRCTACVTC---RDDLVEKTPCAW 82
Db 43 CCLSCPPGPGTYASRLCDSKTNTQCTPCGSDTFTSHNNHLQACLSCNGRCDNQVETRSCNT 102
QY 83 NSSRVCECRPGMFCSTSAVNSCARCFHVSVCAPAGMLVKFPGTAQKNVCEPASPGV-SPA 141
Db 103 THNRICECSPGYYCLLKSGSGCRTCISKTKCGIGYGS-GYTSTGDVICSPCGPGTYSHT 161
QY 142 CASPENCKEPSSGT 155
Db 162 VSSTDKCEPVTSTNT 175
```

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:22:37 ; Search time 17 Seconds
(without alignments)
643.843 Million cell updates/sec

Title: US-09-628-126-2_COPY_19_390

Perfect score: 2080
Sequence: 1 FPQDRPFEDTCHGNPSHYD.....PVALSSTGKPVLDAGPVLFW 372

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2.6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2080	100.0	595	US-08-225-986-2	Sequence 2, Appli
2	2080	100.0	595	US-08-570-923-2	Sequence 2, Appli
3	2080	100.0	595	US-09-580-235-6	Sequence 2, Appli
4	2080	100.0	595	US-09-580-235-6	Sequence 2, Appli
5	2074	99.7	595	US-08-232-087A-2	Sequence 2, Appli
6	2074	99.7	595	US-09-006-353A-9	Sequence 9, Appli
7	2074	99.7	595	US-09-573-986-9	Sequence 9, Appli
8	725	34.9	123	US-08-828-683A-19	Sequence 19, Appli
9	699.5	33.6	122	US-08-232-087A-7	Sequence 7, Appli
10	657.5	31.6	120	US-08-232-087A-8	Sequence 8, Appli
11	429	20.6	74	US-08-866-545-6	Sequence 6, Appli
12	422	20.3	74	US-08-866-545-7	Sequence 7, Appli
13	240.5	11.6	518	US-08-385-229-4	Sequence 4, Appli
14	230	11.1	486	US-08-243-010-1	Sequence 1, Appli
15	225.5	10.8	235	US-09-326-394-4	Sequence 4, Appli
16	225.5	10.8	235	US-09-580-235-2	Sequence 2, Appli
17	225.5	10.8	235	US-09-580-235-8	Sequence 8, Appli
18	225.5	10.8	235	US-09-580-181-2	Sequence 2, Appli
19	225.5	10.8	235	US-09-580-181-8	Sequence 8, Appli
20	225.5	10.8	235	US-09-102-530-2	Sequence 2, Appli
21	225.5	10.8	235	US-09-385-229-2	Sequence 2, Appli
22	225.5	10.8	461	US-08-385-229-2	Sequence 2, Appli
23	225.5	10.8	461	US-08-650-000-2	Sequence 2, Appli
24	225.5	10.8	461	US-09-042-785A-7	Sequence 7, Appli
25	225.5	10.8	461	US-08-477-347-3	Sequence 3, Appli
26	225.5	10.8	461	US-09-006-353A-4	Sequence 4, Appli
27	225.5	10.8	461	US-08-476-862-2	Sequence 2, Appli

28	225.5	10.8	461	US-09-573-986-4	Sequence 4, Appli
29	225.5	10.8	461	US-09-573-986-4	Sequence 4, Appli
30	224.5	10.8	235	US-09-580-235-4	Sequence 4, Appli
31	224.5	10.8	235	US-09-580-235-6	Sequence 6, Appli
32	224.5	10.8	235	US-09-580-181-4	Sequence 4, Appli
33	224.5	10.8	235	US-09-580-181-6	Sequence 6, Appli
34	224.5	10.8	235	US-09-102-530-4	Sequence 4, Appli
35	224.5	10.8	235	US-09-102-530-6	Sequence 6, Appli
36	203.5	9.8	227	US-08-974-022-48	Sequence 48, Appli
37	203.5	9.8	227	US-08-795-445A-48	Sequence 48, Appli
38	203.5	9.8	227	US-08-795-447A-48	Sequence 48, Appli
39	203.5	9.8	227	US-08-974-186-48	Sequence 48, Appli
40	203.5	9.8	227	US-08-795-446B-48	Sequence 48, Appli
41	203.5	9.8	227	US-08-706-945D-134	Sequence 134, Appli
42	200.5	9.6	474	US-08-650-000-4	Sequence 4, Appli
43	200.5	9.6	474	US-09-042-785A-8	Sequence 8, Appli
44	200.5	9.6	474	US-09-573-986-4	Sequence 4, Appli
45	200	9.6	163	US-08-219-237B-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-225-989-2
; Sequence 2, Application US/08225989
; Patent No. 5480981
; GENERAL INFORMATION:
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Smith, Craig A.
; APPLICANT: Armitage, Richard J.
; APPLICANT: Gruss, Hans-Jurgen
; TITLE OF INVENTION: No. 5480981el Cytokine That Binds CD30
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kathryn A. Seese, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple 7.1
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/225,989
; FILING DATE: 12 APRIL, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/966,775
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 907,224
; FILING DATE: 01-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 899,660
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 892,459
; FILING DATE: 02-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 889,717
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2804-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644

```

; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-225-989-2

Query Match 100.0%; Score 2080; DB 1; Length 595;
Best Local Similarity 100.0%; Pred. No. 1.6e-147;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPQDRPFEDTCHGNPSHYDKAVRCCYRCMPGLPTQOCQPORPTDCRKQCEPDYILDEA 60
DB 19 FPQDRPFEDTCHGNPSHYDKAVRCCYRCMPGLPTQOCQPORPTDCRKQCEPDYILDEA 78

QY 61 DRCACVTCSDRDLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHFSVCPAGMIVK 120
DB 79 DRCACVTCSDRDLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHFSVCPAGMIVK 138

QY 121 FPGTAQKNTVCEPASPGVSPACASPENCKEPPSSGTIPQAKPTPVSPATSSASTMPVRGCT 180
DB 139 FPGTAQKNTVCEPASPGVSPACASPENCKEPPSSGTIPQAKPTPVSPATSSASTMPVRGCT 198

QY 181 RLQAEAAKSLTRAPDSPSSVGRPSDDGLSPTQCPGSGDCRKQCEPDYILDEAGRCTA 240
DB 199 RLQAEAAKSLTRAPDSPSSVGRPSDDGLSPTQCPGSGDCRKQCEPDYILDEAGRCTA 258

QY 241 CVSCSRDLDVEKTPCAWNSRRTCECRPGMFCSTSAVNSCARCFHFSVCPAGMIVK 300
DB 259 CVSCSRDLDVEKTPCAWNSRRTCECRPGMFCSTSAVNSCARCFHFSVCPAGMIVK 318

QY 301 EKDTTFEAPPLGTQDCNPTPENGAPASTPTQSLVDSQASKTLPITPSAPVALSSTG 360
DB 319 EKDTTFEAPPLGTQDCNPTPENGAPASTPTQSLVDSQASKTLPITPSAPVALSSTG 378

QY 361 KPVLDAGPVLEW 372
DB 379 KPVLDAGPVLEW 390

RESULT 2
US-08-570-923-2
; Sequence 2, Application US/08570923
; Patent No. 5677430
; GENERAL INFORMATION:
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Smith, Craig A.
; APPLICANT: Armitage, Richard J.
; APPLICANT: Gruss, Hans-Jurgen
; TITLE OF INVENTION: No. 5677430el Cytokine That Binds CD30
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kathryn A. Seese, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple 7.1
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,923
; FILING DATE: 12-DEC-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/225,989
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: US 07/966,775

```

```

; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 907,224
; FILING DATE: 01-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 899,660
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 892,459
; FILING DATE: 02-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 889,717
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2804-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-570-923-2

Query Match 100.0%; Score 2080; DB 1; Length 595;
Best Local Similarity 100.0%; Pred. No. 1.6e-147;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPQDRPFEDTCHGNPSHYDKAVRCCYRCMPGLPTQOCQPORPTDCRKQCEPDYILDEA 60
DB 19 FPQDRPFEDTCHGNPSHYDKAVRCCYRCMPGLPTQOCQPORPTDCRKQCEPDYILDEA 78

QY 61 DRCACVTCSDRDLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHFSVCPAGMIVK 120
DB 79 DRCACVTCSDRDLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHFSVCPAGMIVK 138

QY 121 FPGTAQKNTVCEPASPGVSPACASPENCKEPPSSGTIPQAKPTPVSPATSSASTMPVRGCT 180
DB 139 FPGTAQKNTVCEPASPGVSPACASPENCKEPPSSGTIPQAKPTPVSPATSSASTMPVRGCT 198

QY 181 RLQAEAAKSLTRAPDSPSSVGRPSDDGLSPTQCPGSGDCRKQCEPDYILDEAGRCTA 240
DB 199 RLQAEAAKSLTRAPDSPSSVGRPSDDGLSPTQCPGSGDCRKQCEPDYILDEAGRCTA 258

QY 241 CVSCSRDLDVEKTPCAWNSRRTCECRPGMFCSTSAVNSCARCFHFSVCPAGMIVK 300
DB 259 CVSCSRDLDVEKTPCAWNSRRTCECRPGMFCSTSAVNSCARCFHFSVCPAGMIVK 318

QY 301 EKDTTFEAPPLGTQDCNPTPENGAPASTPTQSLVDSQASKTLPITPSAPVALSSTG 360
DB 319 EKDTTFEAPPLGTQDCNPTPENGAPASTPTQSLVDSQASKTLPITPSAPVALSSTG 378

QY 361 KPVLDAGPVLEW 372
DB 379 KPVLDAGPVLEW 390

RESULT 3
US-08-580-014-2
; Sequence 2, Application US/08580014
; Patent No. 5753203
; GENERAL INFORMATION:
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Smith, Craig A.
; APPLICANT: Armitage, Richard J.
; APPLICANT: Gruss, Hans-Jurgen
; TITLE OF INVENTION: No. 5753203el Cytokine That Binds CD30

```

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Seese, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.1
SOFTWARE: Microsoft Word, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/580,014
FILING DATE: 20-DEC-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/225,989
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: US 07/966,775
FILING DATE: 27-OCT-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 907,224
FILING DATE: 01-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,660
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 892,459
FILING DATE: 02-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 889,717
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2804-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-580-014-2

Query Match 100.0%; Score 2080; DB 1; Length 595;
Best Local Similarity 100.0%; Pred. No. 1.6e-147;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPQDRPFEDTCHGNPSHYDKAVRCCYRCMPGLFPTQOCQORPTDCRKOCEPDYYLDEA 60
DB 19 FPQDRPFEDTCHGNPSHYDKAVRCCYRCMPGLFPTQOCQORPTDCRKOCEPDYYLDEA 78
QY 61 DRCTACVTCRRDLVEKTPCAWNSRRTCECRPGMICATSATNSCARCVYPICAGETVTKPDMA 120
DB 79 DRCTACVTCRRDLVEKTPCAWNSRRTCECRPGMICATSATNSCARCVYPICAGETVTKPDMA 138
QY 121 FPGTAQKNTVCEPASGVSPACAPENCKEPPSSGTTIPQAKPTVPSPATSSATMPVRGCT 180
DB 139 FPGTAQKNTVCEPASGVSPACAPENCKEPPSSGTTIPQAKPTVPSPATSSATMPVRGCT 198
QY 181 RLQAQEAASKLTPAPSPSSVGRPSDDPGLSPTQCPGEGSGDKCRKOCEPDYYLDEAGRCTA 240
DB 199 RLQAQEAASKLTPAPSPSSVGRPSDDPGLSPTQCPGEGSGDKCRKOCEPDYYLDEAGRCTA 258
QY 241 CVSCSRDDLVEKTPCAWNSRRTCECRPGMICATSATNSCARCVYPICAGETVTKPDMA 300

DB 259 CVSCSRDDLVEKTPCAWNSRRTCECRPGMICATSATNSCARCVYPICAGETVTKPDMA 318
QY 301 EKDTTFEAPPLGTQPCDNCPTPENGEAPASTPTQSLLDVDSQASKTLPIPTSPAPVALSSTG 360
DB 319 EKDTTFEAPPLGTQPCDNCPTPENGEAPASTPTQSLLDVDSQASKTLPIPTSPAPVALSSTG 378
QY 361 KPVLDAGPVLFW 372
DB 379 KPVLDAGPVLFW 390
RESULT 4
US-09-079-785-2
Sequence 2, Application US/09079785
Patent No. 6143869
GENERAL INFORMATION:
APPLICANT: Goodwin, Raymond G.
APPLICANT: Smith, Craig A.
APPLICANT: Armitage, Richard J.
APPLICANT: Gruss, Hans-Jurgen
TITLE OF INVENTION: No. 6143869el Cytokine That Binds CD30
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Seese, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.1
SOFTWARE: Microsoft Word, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,785
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/225,989
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: US 07/966,775
FILING DATE: 27-OCT-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 907,224
FILING DATE: 01-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,660
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 892,459
FILING DATE: 02-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 889,717
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2804-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-079-785-2
Query Match 100.0%; Score 2080; DB 4; Length 595;

```

Best Local Similarity   100.0%; Pred. No. 1.6e-147;
Matches 372; Conservative    0; Mismatches    0; Indels      0; Gaps      0;

QY      1 FPQDRPFEDTCHGNPSHYIDKAVRRCCYRCPMGLFPTQCQPQRPTDCRKQCEPDYYLDEA  60
       |||||||
DB     19 FPQDRPFEDTCHGNPSHYIDKAVRRCCYRCPMGLFPTQCQPQRPTDCRKQCEPDYYLDEA  78

QY     61 DRCTACVTCSRDDLVEKTTPCAWNSRVCCECRPGMFCSTSAVNSCARCFHFHSCVPAGMIVK 120
       |||||||
DB     79 DRCTACVTCSRDDLVEKTTPCAWNSRVCCECRPGMFCSTSAVNSCARCFHFHSCVPAGMIVK 138

QY    121 FPGTAQNTVCEPASPGVSPACASPENCKEPSSGTIPQAKETPVSPATSSASTHPVRGGT 180
       |||||||
DB   139 FPGTAQNTVCEPASPGVSPACASPENCKEPSSGTIPQAKETPVSPATSSASTHPVRGGT 198

QY    181 RLAQEAASKLTRAPDSPSSVGRPSDDPGLSPTQCPPEGSGDCKRQCEPDYYLDEAGRCTA  240
       |||||||
DB   199 RLAQEAASKLTRAPDSPSSVGRPSDDPGLSPTQCPPEGSGDCKRQCEPDYYLDEAGRCTA  258

QY    241 CVCSRDLDVEKTTPCAWNSRRTCECRPGMICATSATNSCARCVYPICAGETVTTKPDMA  300
       |||||||
DB   259 CVCSRDLDVEKTTPCAWNSRRTCECRPGMICATSATNSCARCVYPICAGETVTTKPDMA  318

QY    301 EKDTTFEAPPLGTOPDCNPTPENGEAFASPTOSLLVDISOASKTLPIPTAPVALSSTG  360
       |||||||
DB   319 EKDTTFEAPPLGTOPDCNPTPENGEAFASPTOSLLVDISOASKTLPIPTAPVALSSTG  378

QY    361 KPVLDAGPVLFW 372
       |||||||
DB   379 KPVLDAGPVLFW 390

RESULT 5
US-08-232-087A-2
; Sequence 2, Application US/08232087A
; Patent No. 5866372
; GENERAL INFORMATION:
; APPLICANT: Stein, Harald
; APPLICANT: D rkop, Horst
; APPLICANT: Latza, Ute
; TITLE OF INVENTION: Lymphoid CD30-Antigen
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,087A
; FILING DATE: 08-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 756-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-232-087A-2

```

(GENE-) GENE LOGIC INC.
Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
WPI; 2002-435328/46.
Detecting granulocyte activation by detecting differential expression
of genes associated with granulocyte activation, which serves as
diagnostic markers that is useful for monitoring disease states and
drug toxicity.
Claim 1; SEQ ID No 387; 114pp; English.
The invention relates to detecting (M1) granulocyte (GC) activation
(GCA), by detecting the level of expression of gene(s) (Gs) identified by
DNA chip analysis as given in the specification, and comparing
the expression level to an expression level in an unactivated
GC, where differential expression of Gs is indicative of GCA.
Also included are modulation of at least one gene in Gs; (2) screening (M3)
for an agent capable of modulating GCA or an inflammation (especially
chronic) in a tissue, an allergic response in a subject, exposure of a
subject to a pathogen or sterile inflammatory disease using the
gene expression profile; (3) detecting (M4) an inflammation (especially
chronic) in a tissue, an allergic response in a subject, exposure of a
subject to a pathogen or sterile inflammatory disease, by detecting the
level of expression in a sample of the tissue of gene(s) from Gs, where
the level of expression of the gene is indicative of inflammation;
(4) treating (M5) an inflammation (especially chronic) or in a tissue,
an allergic response in a subject, exposure of a subject to a pathogen
or sterile inflammatory disease, by contacting a tissue having
inflammation with an agent that modulates the expression of gene(s)
from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
GCA preferably in an inflammation in a tissue; M4 is useful for
detecting an inflammation (especially chronic) in a tissue, an allergic
response in a subject, exposure of a subject to a pathogen or sterile
inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
refusion injury, ARDS, adult respiratory distress syndrome,
inflammatory bowel disease, Crohn's disease, ulcerative colitis,
periodontal disease; also bacterial infection, viral infection,
parasitic infection, protozoal infection, fungal infection and M5 is
useful for treating one of the above conditions. The present
sequence represents a gene differentially expressed in granulocytes.
Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic
format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.
Sequence 1906 BP; 559 A; 447 C; 438 G; 462 T; 0 other;
Query Match 100.0%; Score 648; DB 24; Length 1906;
Best Local Similarity 100.0%; Pred. No. 9.5e-198;
Matches 648; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCATGTCCGGCGGCTCCGTGGCCAGCCACTGGGACACGACGCGCCACTATTTC 60
DB 172 ATGCATGTCCGGCGGCTCCGTGGCCAGCCACTGGGACACGACGCGCCACTATTTC 231
QY 61 TATTTGACCACAGCCACTCTGGCTGTGCTTTCACGGTGGCCACTATTATGGTG 120
DB 232 TATTTGACCACAGCCACTCTGGCTGTGCTTTCACGGTGGCCACTATTATGGTG 291
QY 121 TTGGTGGTTTCAGAGGACGACTCCATTCCTCCAACTACCTGACACGCTCCCTCAAGGA 180
DB 292 TTGGTGGTTTCAGAGGACGACTCCATTCCTCCAACTACCTGACACGCTCCCTCAAGGA 351
QY 181 GGAATTTGTCAGAGACCTCTTATGTATCTCGAAAGAGCTCCATCAAGAGTCATGG 240
DB 352 GGAATTTGTCAGAGACCTCTTATGTATCTCGAAAGAGCTCCATCAAGAGTCATGG 411
QY 241 GCCTACCTCCAAAGTGCAAGCATCTTAACAAACCAAGTTGTCTTGGACAAAGATGGC 300

DB 412 GCCTACCTCCAAAGTGCAAGCATCTTAACAAACCAAGTTGTCTTGGACAAAGATGGC 471
QY 301 ATTCTCCATGGAGTCAGATATCAGGATGGATCTGGTGATCCAATTCCTCGTTTGTAC 360
DB 472 ATTCTCCATGGAGTCAGATATCAGGATGGATCTGGTGATCCAATTCCTCGTTTGTAC 531
QY 361 TTCTATCATTTTCCCAACTGCAGTTTCTTGTACAAATGCCCAAAATAATTTCTGTCGATCTGAAG 420
DB 532 TTCTATCATTTTCCCAACTGCAGTTTCTTGTACAAATGCCCAAAATAATTTCTGTCGATCTGAAG 591
QY 421 TTGGAGCTTTCATCAACAAGCATATCAAAAAAGAGCCCTGGTGACAGTGTGTGAGTCT 480
DB 592 TTGGAGCTTTCATCAACAAGCATATCAAAAAAGAGCCCTGGTGACAGTGTGTGAGTCT 651
QY 481 GGAATGCAACGAAACAGCTATACCAAGATCTCTCTCAATTTCTGCTGGATTACCTGCAG 540
DB 652 GGAATGCAACGAAACAGCTATACCAAGATCTCTCTCAATTTCTGCTGGATTACCTGCAG 711
QY 541 GTCAACACACCATATCATCTAGTCAATGTGTGATACATTCAGTACATAGATACAAGACCTTT 600
DB 712 GTCAACACACCATATCATCTAGTCAATGTGTGATACATTCAGTACATAGATACAAGACCTTT 771
QY 601 CCTCTTGAGAAATGTCTTGTCCATCTTCTTATACAGTAATTCAGACTGA 648
DB 772 CCTCTTGAGAAATGTCTTGTCCATCTTCTTATACAGTAATTCAGACTGA 819
RESULT 6
AAQ53537
ID AAQ53537 standard; CDNA; 720 BP.
XX AC AAQ53537;
XX DT 19-JUN-1994 (first entry)
XX DE Sequence of the coding region of a murine CD30-L cDNA clone
DE encoding additional N-terminal amino acids.
XX KW Hodgkin's disease; lymphoma; surface antigen; cytokine;
XX CD30 ligand; CD30-L; TNF; NGF; ss.
XX OS Acromys cahirinus.
XX PH Key Location/Qualifiers
XX CDS 1..720
XX FT /*tag= a
XX PN WO9324135-A.
XX PD 09-DEC-1993.
XX XX 25-MAY-1993; 93WO-US04926.
XX PF 26-MAY-1992; 92US-0889717.
XX PR 02-JUN-1992; 92US-0892459.
XX PR 13-JUN-1992; 92US-0899660.
XX PR 01-JUL-1992; 92US-0907224.
XX PR 27-OCT-1992; 92US-0966775.
XX PA (IMMV) IMMUNEX CORP.
XX PI Armitage RJ, Goodwin RG, Smith CA;
XX WPI; 1993-405417/50.
XX DR P-PSDB; AAR45008.
XX XX New cytokine, CD30-L, which binds CD30 - used for developing
PT prods. for diagnosis, detection, purifications, research and
PT therapy
XX PS Claim 1; Figure 6a; 59pp; English.

CC CD30-L is a ligand for CD30, the 120kd surface antigen widely used
CC as a clinical marker for Hodgkin's lymphoma and related haematologic
CC malignancies. CD30 is a member of the TNF/NGF superfamily. The CD30-
CC L and other derived prods. can be used for elucidating the roles
CC that CD30 and CD30-L may play in the immune system and for diagnosis
CC and therapy. It can be isolated as follows. A cDNA library prep'd from
CC the murine helper T-cell line 7B9 is screened with a CD30/Fc fusion
CC protein labelled with (125)I to obtain cDNA encoding murine CD30-L
CC (AAQ53535). This cDNA can then be used as a probe to screen a human PBL
CC cDNA library to obtain cDNA encoding human CD30-L (AAQ53536). An
CC anchored PCR technique was employed to isolate CD30-L human and murine
CC clones containing an additional 19 N-terminal amino acid sequence
CC (AAQ53537, AAQ53538).

XX Sequence 720 BP; 191 A; 196 C; 161 G; 172 T; 0 other;

Query Match 62.3%; Score 403.8; DB 14; Length 720;
Best Local Similarity 79.1%; Pred. No. 2.3e-119; Indels 3; Gaps 1;
Matches 493; Conservative 0; Mismatches 127;

QY 26 CCAGCCACTGGGGACCCAGCGCAGCTATTTCTATTGACCAGCCACTCTGGCTC 85

Db 101 CGAGCCCTGGAGAGACACAGTCGCGACTTCTTACCTCAGCACCCCGCACTGG--- 157

QY 86 TGTGCTTGTCTTACGCTGGCCACTATTATGGTGTGTGCTGTTTCAGAGCGGACTCCA 145

Db 158 TGTGCTTGTGTGCGAGTGGCGATCATTTCTGGTACTGTCCAGAAAAAGGACTCCA 217

QY 146 TTCCCAACTCACCTGACAACTGCCCCCTCAAGAGGAGAAATTTGCTCAGAGACCTCTTAT 205

Db 218 CTCCAAATACAACTGAGAGGCGCCCTTAAAGGAGGAAATTTGCTCAGAGGATCTTCT 277

QY 206 GTATCCTGAAAAGAGCTCCATTCAGAAAGTCAATGGCCCTACCTCCAAAGTGGCAAGCATC 265

Db 278 GTACCTGAAAAGTACTCCATCCAGAGAGTCAATGGCCCTACCTCCAAAGTGTCAAGCATC 337

QY 266 TAAACAAAACCAAGTTGCTTTGGAAACAAAGATGGCATTTCTCCATGGAGTCAGATATCAGG 325

Db 338 TCAACAAATACCAAACTGTGTAAGACGAAGATGGCACCATCCACGACTCATATACCAAG 397

QY 326 ATGGAAATCGGTGATCAAAATCCCTGGTTGTACTTCAATCATATTTGCCAATGCAATTC 385

Db 398 ACGGGAACCTGATAGTCCAAATCCCTGGCTTGTACTTCAATCATATTTGCCAATGCAATTC 457

QY 386 TTGTACAATGCCAAATATTTCTGTCGATCTGAAGTGGAGCTTCTCAATCAACAAGCAT 445

Db 458 TCGTCACTGTCAAAATCAATCTGTGGACCTGACATTTGCAGCTCTCATCAATTCACAGA 517

QY 446 TCAAAAACAGCCCTGGTGACAGTGTGTGAGTCTGGATGCAAAACGAAACATATACC 505

Db 518 TCAAAAACAGAGGTTGTAACAGTGTGTGAGTCTGGAGTTCAGAGTAAGAACAATCTACC 577

QY 506 AGAATCTCTCAATTTCTGTTGATTTACCTGCAGGTCAACACCACCATATCAGTCAATG 565

Db 578 AGAATCTCTCAGTTTGTGTCATTTCTTACAGGTCAACCTCACTACCATATCAGTCAGG 637

QY 566 TGGATACATTCAGTACATAGATACAGCACTTTCTCTTTCAGATGCTGTTGTCATCT 625

Db 638 TGGAAATTTTCAGTATGTGGTACAAACACTTTCCCTCTTGTATATGTCATCCGCT 697

QY 626 TCTTATACAGTAATTCAGACTGA 648

Db 698 TCTTATATAGTACGACTGA 720

RESULT 7

AAQ53535

XX AAQ53535 standard; cDNA; 663 BP.

XX AAQ53535;

XX 19-JUN-1994 (first entry)

XX

DE Sequence of the coding region of a murine CD30-L cDNA clone.

XX Hodgkin's disease; lymphoma; surface antigen; cytokine;

KW CD30 ligand; CD30-L; TNF; NGF; ss.

XX Acomys cahirinus.

XX Key Location/Qualifiers

FT CDS 1..663

XX /*tag= a

PN W09324135-A.

XX 09-DEC-1993.

XX 25-MAY-1993; 93WO-US04926.

XX 26-MAY-1992; 92US-0889717.

PR 02-JUN-1992; 92US-0892459.

PR 15-JUN-1992; 92US-0899660.

PR 01-JUL-1992; 92US-0907224.

PR 27-OCT-1992; 92US-0966775.

XX (IMMV) IMMUNEX CORP.

XX Armitage RJ, Goodwin RG, Smith CA;

PI WPI: 1993-405417/50.

DR P-PSDB; AAR45006.

XX New cytokine, CD30-L, which binds CD30 - used for developing

PT prods. for diagnosis, detection, purifications, research and

PS therapy

XX Claim 1; Figure 3a; 59pp; English.

XX CD30-L is a ligand for CD30, the 120kd surface antigen widely used
CC as a clinical marker for Hodgkin's lymphoma and related haematologic
CC malignancies. CD30 is a member of the TNF/NGF superfamily. The CD30-
CC L and other derived prods. can be used for elucidating the roles
CC that CD30 and CD30-L may play in the immune system and for diagnosis
CC and therapy. It can be isolated as follows. A cDNA library prep'd from
CC the murine helper T-cell line 7B9 is screened with a CD30/Fc fusion
CC protein labelled with (125)I to obtain cDNA encoding murine CD30-L.
CC This cDNA can then be used as a probe to screen a human PBL cDNA
CC library to obtain cDNA encoding human CD30-L.

XX Sequence 663 BP; 180 A; 175 C; 144 G; 164 T; 0 other;

Query Match 62.1%; Score 402.2; DB 14; Length 663;

Best Local Similarity 79.0%; Pred. No. 7.2e-119;

Matches 492; Conservative 0; Mismatches 128; Indels 3; Gaps 1;

QY 26 CCAGCCACTGGGACCCAGCGCAGCTATTTCTATTGACCAGCCACTCTGGCTC 85

Db 44 CGAGCCCTGGAGAGACACAAAGTCGCGACTTCTTACCTCAGCACCCCGCACTGG--- 100

QY 86 TGTGCTTGTCTTACGCTGGGCCACTATTATGGTGTGCTGTTTCAGAGCGGACTCCA 145

Db 101 TGTGCTTGTGTGCGAGTGGCGATCATTTCTGGTACTGTGTAGTCCAGAAAAGACTCCA 160

QY 146 TTCCCAACTCACCTGACAACTGCCCCCTCAAGAGGAGAAATTTGCTCAGAGACCTCTTAT 205

Db 161 CTCCAAATACAACTGAGAGGCGCCCTTAAAGAGGAAATTTGCTCAGAGATCTTCTTCT 220

QY 206 GTATCTCTGAAAAGAGCTCCATTCAGAAAGTCATGGGCTACCTCCAAAGTGCAGGATC 265

Db 221 GTACCTCTGAAAAGTACTCCATCCCAAGAGTCATGGGCTACCTCCCAAGTGTCAAGCATC 280

QY 266 TAAACAAAACCAAGTTGCTTTGGAACAAAGATGGCATTTCTCCATGGAGTCAGATATCAGG 325

Db 281 TCAACAAATACCAAACTGTCTATGGAAGAGATGCGACCATCCCGGACTCATATACCAAG 340

QY 326 ATGGGAATCGTGATCCAAATCCCTGGTTTGTACTTTCATCATTTGCCAACTGCAGTTTC 385
 Db 341 ACGGAAACCTGATAGTCCCAATTCCTGGCTTGTACTTTCATCGTTTGGCCAAGTCGATTC 400
 QY 386 TTGTACAATGCCAAATTAATCTGTCGATCTCAAGTTGGAGCTTCTCATCAACAAGCATA 445
 Db 401 TCGTGCAGTCTCAATCATCTGTGGACCTGCACATTCGAGCTCCTCATCAATTCCAAGA 460
 QY 446 TCAAAAACAGCCCTGGTGACAGTGTGTGAGTCTGGAATGCAAAACGAAACAGTATACC 505
 Db 461 TCAAAAACAGAGCTGTGTAAACAGTGTGTGAGTCTGGAGTTCAGAGTAAAGACATCTACC 520
 QY 506 AGAATCTCTCAATCTTGTGGAATACCTGCAGGTCAACACCACCATATACAGTCAATG 565
 Db 521 AGAATCTCTCAGTTTGTGCTGCAATTAATACAGGTCAACCTCTACCATATCAGTCAGGG 580
 QY 566 TGGATACATTCAGTACATAGATCAACAGCCTTTCCTCTTGAGAAATGTTGTCCATCT 625
 Db 581 TGGATAATTTCCAGTATGTGGATACAAACACTTCCCTCTTGATAATGTGCTATCCGTCT 640
 QY 626 TCTTATACAGTAATTCAGACTGA 648
 Db 641 TCTTATATAGTACGACTGA 663

RESULT 8
 AAV42207
 ID AAV42207 standard; DNA; 720 BP.
 XX
 AC AAV42207;
 DT 23-SEP-1998 (first entry)
 XX
 DE CD30 ligand gene used in the course of the invention.
 XX
 KW CD30 ligand; alteration; immunoreactivity; human cell;
 KW accessory molecule ligand; AML; gene therapy; treatment; neoplasia;
 KW autoimmune disorder; rheumatoid arthritis; vaccine; ss.
 XX
 OS Mus sp.
 XX
 PN W09826061-A2.
 XX
 PD 18-JUN-1998.
 XX
 PF 08-DEC-1997; 97WO-US22740.
 XX
 PR 01-DEC-1997; 97US-0982272.
 PR 09-DEC-1996; 96US-0032145.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Cantwell M, Kipps TJ, Sharma S;
 DR WPI; 1998-348521/30.
 XX
 PT Vectors containing accessory molecule ligand genes - used for
 PT altering immunoreactivity of cells, particularly for treatment of
 PT neoplasia or autoimmune disorders, e.g. rheumatoid arthritis
 XX
 PS Disclosure; Page 125; 167pp; English.
 XX
 CC The present sequence represents a CD30 ligand gene. The sequence is
 CC used to exemplify the method of the invention. The specification
 CC describes a method for altering the immunoreactivity of human cells
 CC which comprises introducing a gene encoding an accessory molecule
 CC ligand (AML) into the cells so that the AML is expressed on the surface
 CC of the cells. Vectors containing the AML genes can be used in the
 CC therapy for treating neoplasia or autoimmune disorders such as rheumatoid
 CC arthritis. They can also be used for vaccination to produce immunity
 CC against a virus cell, bacteria, protein, fungus or neoplasia.
 XX
 SQ Sequence 720 BP; 191 A; 195 C; 162 G; 172 T; 0 other;

Query Match 62.1%; Score 402.2; DB 19; Length 720;
 Best Local Similarity 79.0%; Pred. No. 7.5e-119;
 Matches 492; Conservative 0; Mismatches 128; Indels 3; Gaps 1;
 QY 26 CCAGCCACCTGGGGACCCAGAGCCGAGCTATTTCATTGACACAGCACCTCTGGCTC 85
 Db 101 CGAGGCCCTGGAGAAGCACAAGTCGAGCTACTTCTACCTCAGCACCCGCCACTGG-- 157
 QY 86 TGTGCCCTCTCTCACGGTGGCCACTATTATGTTGTTGTCGTTTCAGAGACGACATCCA 145
 Db 158 TGTGCCCTTGTGGCAGTGGCGATCATCTGTTACTGGTAGTCCAGAAAAGAGACTCCA 217
 QY 146 TTCCCAACTCACTCACAACGTCCTCCCTCAAAGGAGAAATTCCTCAGAAGACTCTTAT 205
 Db 218 CTCCAATACAACTGAGAGGCCCTTAAAGAGGAATTCCTCAGAGGATCTCTTCT 277
 QY 206 GTATCTGAAAAGAGCTCCATTCAGAAGTCATGGGCTACCTCCAAAGTGGCAAGATC 265
 Db 278 GTACCTGAAAAGTACTCCATCCCAAGATCATGTTGGGCTACCTCCAAGTGTCAAAGCATC 337
 QY 266 TAAACAAAACCAAGTTGTTCTTGGAAACAAGATGGCATTCTCCATGGAGTCAGATATCAGG 325
 Db 338 TCAACAATACCAAACTGTCATGGAGCAAGATGGCCACCATCCAGGACTCATATACCAGG 397
 QY 326 ATGGGAATCTGGTGATCCAAATCCCTGGTTTGTACTTTCATCATTTGCCAACTGCAGTTTC 385
 Db 398 ACGGNAACCTGATAGTCCAAATCCCTGGCTTGTACTTTCATCTGTTGCCAACTGCAGTTTC 457
 QY 386 TTGTACAATGCCCAATAATTCCTGTCGATCTGAAGTTGGAGCTTCTCATCAACAAGCATA 445
 Db 458 TCGTGCAGTGTCAAAATCATCTGTGACCTGACATTCGAGCTCTCATCAATTCGAAGA 517
 QY 446 TCAAAAACAGGCCCTGGTGACAGTGTGAGTCTGGAATGCAAAACGAAACAGTATACC 505
 Db 518 TCAAAAACAGCAGCTGTGGTAAACAGTGTGAGTCTGGAGTTCAGAGTAAAGAACATCTACC 577
 QY 506 AGAATCTCTCAATTCCTTGGATTACTCTGAGGTCAACACCAATATCACTCAATG 565
 Db 578 AGAATCTCTCAGTTTTTGTGCTACTTACAGGTCAACTTACCATATCAGTACAGG 637
 QY 566 TGGATACATCCAGTACATAGATACAAAGCACTTTCCTCTTGAGAAATGTGTTGCCATCT 625
 Db 638 TGGATAATTTCCAGTATGTGGATACAAACACTTTCCTCTTGATAATGTGCTATCCGTCT 697
 QY 626 TCTTATACAGTAATTCAGACTGA 648
 Db 698 TCTTATATAGTACGCTCAGACTGA 720

RESULT 9
 ABK12323
 ID ABK12323 standard; DNA; 720 BP.
 XX
 AC ABK12323;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE Mouse CD30-ligand (CD30L) gene sequence.
 XX
 KW Mouse; CD30-ligand; CD30L; CD30; IL; interleukin; IL-lambda; IL-lbeta;
 KW IL-RL; signal transduction; autoimmune condition; multiple sclerosis;
 KW chronic inflammatory condition; systemic sclerosis; Fisher syndrome;
 KW inflammatory demyelinating polyneuropathy; motor axonal neuropathy;
 KW motor sensory axonal neuropathy; systemic lupus erythematosus;
 KW vulgaris; rheumatic disorder; endocrine system disorder; allergy;
 KW gastrointestinal system disorder; genitourinary system disorder;
 KW haematologic disorder; hereditary condition; liver disorder;
 KW lung disease; transplantation disorder; degenerative disease;
 KW skin disorder; mucous membrane disorder; sarcoidosis; arteritis;
 KW multicentric reticulohistiocytosis; Wegener's granulomatosis; vasculitis;
 KW arthritic condition; TNFalpha inhibitor; gene; ds.


```
RESULT 14
AAK49336
ID AAK49336 standard; DNA; 240 BP.
XX
XX
AC AAK49336;
XX
XX
DT 06-NOV-2001 (first entry)
XX
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 23893.
XX
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200157276-A2.
XX
XX
PD 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX
DR WPI; 2001-488900/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX
XX
XX
Example 4; SEQ ID NO: 23893; 658pp + Sequence Listing; English.
PS
XX
The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
XX
Sequence 240 BP; 73 A; 50 C; 52 G; 65 T; 0 other;
SQ
Query Match 37.0%; Score 240; DB 22; Length 240;
Best Local Similarity 100.0%; Pred. No. 6.9e-67;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 254 TGGCAACGATCTAAACAAACCAAGTTGCTCTGGCAACAAGATGCGATCTCCATGGAG 313
Db 1 TGGCAACGATCTAAACAAACCAAGTTGCTCTGGCAACAAGATGCGATCTCCATGGAG 60
QY 314 TCAGATATCAGGATGGGAATCGTGTGATCCCAATTCCTGGTTGTACTTCATCATTTGCC 373
Db 61 TCAGATATCAGGATGGGAATCGTGTGATCCCAATTCCTGGTTGTACTTCATCATTTGCC 120
QY 374 AACTGCAGTTCTTGTACAAATGCCCAATAATTCGTGCATCTGAAGTTGGAGCTTCTCA 433
Db 121 AACTGCAGTTCTTGTACAAATGCCCAATAATTCGTGCATCTGAAGTTGGAGCTTCTCA 180
QY 434 TCACAACGATATCAAAAACAGGCCCTGGTGACAGTGTGTGAGTCTGGAATGCAACGA 493
Db 181 TCACAACGATATCAAAAACAGGCCCTGGTGACAGTGTGTGAGTCTGGAATGCAACGA 240
RESULT 15
```

```
AAS46771
ID AAS46771 standard; DNA; 17280 BP.
XX
XX
AC AAS46771;
XX
XX
DT 18-DEC-2001 (first entry)
XX
XX
DE Tumour suppressor gene derived chemically modified sequence #495.
XX
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200168912-A2.
XX
XX
PD 20-SEP-2001.
XX
XX
PF 15-MAR-2001; 2001WO-EP02955.
XX
XX
PR 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX
PA (EPIC-) EPIGENOMICS AG.
XX
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
XX
DR WPI; 2001-602752/68.
XX
XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer -
XX
XX
XX
Claim 1; SEQ ID No 495; 27pp; English.
PS
XX
The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC and sequences complementary to (Ss). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.
CC
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
Sequence 17280 BP; 5287 A; 151 C; 3639 G; 8203 T; 0 other;
SQ
```

```
Query Match 13.3%; Score 86.2; DB 22; Length 17280;
Best Local Similarity 76.3%; Pred. No. 2.1e-16;
Matches 106; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 1 ATCATGTGCGCGGGCTCCGTGGCCAGCACCTCGGGACACGACGCGCGAGCTATTTC 60
Db 5172 AUGTATGTGCGCGGGCTTCGTGGTTAGTTATTGGGATTACGAGTCGTAGTATTTT 5231
```

